ÉTICA INTUICIONISTA

Michael Huemer

ΔCRΔCIΔ

ÉTICA INTUICIONISTA

Michael Huemer

Δ C R Δ C I Δ
EDITORIAL

Í N D I C E

Índice analítico 4

Prefacio 21

1 Introducción 27

PARTE I: Teorías metaéticas alternativas 48

2 El no cognitivismo 49

3 Subjetivismo 101

4 El reduccionismo 131

PARTE II: Intuicionismo ético 183

5 El conocimiento moral 184

6 Desacuerdo y error 234

7 Razones prácticas 278

8 Otras objeciones 350

9 Conclusión 393

Notas 437

Referencias 439

Índice Analítico

1 INTRODUCCIÓN 27

1.1 El campo de la metaética 27

La metaética aborda cuestiones sobre la naturaleza de los enunciados y juicios evaluativos, incluyendo preguntas sobre el significado del discurso evaluativo, nuestro conocimiento del valor, la objetividad del valor y cómo los juicios de valor proporcionan razones para la acción.

1.2 ¿Qué es la objetividad? 28

Una propiedad objetiva de una cosa es aquella que no depende constitutivamente de las actitudes de los observadores hacia esa cosa.

1.3 Cinco teorías metaéticas 31

Existen exactamente cinco teorías metaéticas: el no cognitivismo, el subjetivismo, el nihilismo, el naturalismo y el intuicionismo.

1.4 Una taxonomía alternativa de los puntos de vista metaéticos 36

Las teorías metaéticas pueden dividirse en monistas y dualistas. Las teorías monistas pueden dividirse a su vez en teorías reduccionistas y eliminativistas; pero estas teorías sólo se diferencian entre sí semánticamente. Sólo el intuicionismo es dualista y difiere fundamentalmente de todas las demás teorías.

1.5 Un intuicionismo racionalista 39

Defenderé una forma de intuicionismo según la cual términos como "bueno" se refieren a propiedades de valor objetivas e irreductibles, que conocemos sobre la base de la intuición racional, y nuestros juicios evaluativos nos dan razones para la acción independientemente de

nuestros deseos.

1.6. Supuestos previos 43
Asumo que existen hechos y conocimientos objetivos fuera del ámbito de la ética. Argumentaré que la ética no es diferente.

PARTE I: TEORÍAS METAÉTICAS ALTERNATIVAS 48

2 EL NO COGNITIVISMO 49

2.1 El no cognitivismo clásico 49
Los no cognitivistas sostienen que los enunciados evaluativos no afirman proposiciones, sino que son más bien imperativos o expresiones de emoción.

2.2 ¿Cómo podemos saber si el cognitivismo es verdadero? 51
El no cognitivismo debe comprobarse empíricamente, examinando cómo utilizamos el lenguaje evaluativo.

2.3 La evidencia lingüística del cognitivismo 53
En el lenguaje ordinario, los enunciados evaluativos se comportan de todas las maneras discernibles como enunciados cognitivos. Los no cognitivistas tienen dificultades para dar sentido a muchos enunciados ordinarios que incluyen términos evaluativos.

**2.4 Hare sobre la verdad moral* 62
Hare intenta sin éxito explicar por qué podemos llamar "verdaderos" y "falsos" a los enunciados evaluativos.

**2.5 Los mundos fácticos-normativos de Gibbard* 68
La noción de Gibbard de mundos factuales-normativos no ofrece ninguna solución al problema de Frege-Geach.

**2.6 Las soluciones de Blackburn* 74
La solución de Blackburn al problema de Frege-Geach se basa en

interpretaciones erróneas del lenguaje y no aborda varios problemas relacionados con el no cognitivismo.

**2.7 El no descriptivismo de Timmons* 81
La teoría de Timmons no puede dar cuenta del error moral y no puede explicar la distinción entre realismo y antirrealismo.

2.8 La evidencia introspectiva 90
Introspectivamente, los juicios morales parecen igual que las creencias, y no como las emociones, los deseos u otros estados no cognitivos.

3 EL SUBJETIVISMO 101

3.1 ¿Qué es el subjetivismo? 101
Los subjetivistas sostienen que los enunciados evaluativos informan de las actitudes de los observadores hacia los objetos de la evaluación. Las diferentes formas de subjetivismo invocan, respectivamente, las actitudes de los individuos, de la sociedad, de Dios o de observadores ideales hipotéticos.

3.2 Subjetivismo individualista 102
El subjetivismo individualista implica que los enunciados morales nazis son verdaderos, que el desacuerdo moral es imposible, que yo soy moralmente infalible y que las actitudes arbitrarias generan obligaciones.

3.3 El relativismo cultural 105
El relativismo cultural adolece de problemas análogos.

3.4 La teoría del mandato divino 110
La teoría del mandato divino de la ética adolece de algunos de los mismos problemas. Además, se cuestiona si Dios existe y cómo podemos saber lo que quiere en caso de que lo haga.

**3.5 La teoría del observador ideal* 119
Si tenemos cuidado de que la teoría no sea circular, la teoría del observador ideal se enfrenta a problemas análogos a los que tienen el

subjetivismo individual y el relativismo cultural.

3.6 La falacia subjetivista 124
Los argumentos más comunes a favor del subjetivismo se basan en confusiones transparentes entre creencia y verdad, y entre lo que causa una creencia y el contenido de la misma.

4 EL REDUCCIONISMO 131

4.1 ¿Qué es el reduccionismo? 131
Los reduccionistas creen (i) que lo que significa que una cosa es buena puede explicarse utilizando expresiones no evaluativos, y (ii) que conocemos las verdades morales sobre la base de la observación.

4.2 Reduccionistas analíticos 133
Los reduccionistas analíticos creen que alguna expresión no valorativa es sinónimo de "bueno". Esto es refutado por el argumento de la pregunta abierta de G. E. Moore.

4.3 La Guillotina de Hume 140

4.3.1 La Ley de Hume: una premisa inicial 140
Es imposible deducir válidamente un enunciado evaluativo a partir de premisas no evaluativas.

**4.3.2 El desafío de Searle* 144
El intento de contraejemplo de Searle fracasa debido al equívoco.

**4.3.3 La impugnación de Geach* 146
El intento de contraejemplo de Geach fracasa porque es inválido y una de sus premisas es evaluativa.

**4.3.4 La impugnación de Prior* 149
El contraejemplo de Prior carece de interés porque no puede proporcionar un modelo plausible de cómo se adquiere el conocimiento ético típico.

*4.3.5 *La prueba de Karmo* 150
Karmo ha demostrado en general que no hay ninguna derivación sólida de una proposición evaluativa no trivial a partir de premisas no evaluativas, donde una proposición evaluativa es aquella cuya verdad, una vez fijados todos los hechos naturales, depende de qué sistema de valores sea correcto.

4.4 Reduccionistas sintéticos 156
Los reduccionistas sintéticos sostienen que, aunque el significado de "bueno" no puede darse utilizando expresiones no evaluativas, se puede explicar qué es la bondad utilizando expresiones no evaluativas.

4.4.1 ¿Pueden conocerse los hechos morales mediante observación? 147
Incluso si las propiedades morales son reducibles, sería falaz deducir que podemos conocer las verdades morales por observación. No podemos observar que una cosa es buena, porque no hay ninguna forma distintiva de que las cosas buenas se vean, suenen, huelan, sepan o sientan.

4.4.2 ¿Pueden conocerse los hechos morales por inferencia a la mejor explicación? 162
Aunque algunos hechos morales sean explicativos, no podemos conocer las verdades morales por inferencia a la mejor explicación , porque los hechos morales no explican ninguna observación que no pueda ser explicada también por hechos no morales.

4.4.3 ¿Se pueden probar las afirmaciones morales? 167
Las teorías morales no generan ninguna predicción comprobable sin basarse en postulados *ad hoc* o en la suposición de que los seres conscientes tienen algún acceso independiente a las verdades morales.

4.4.4 El poder unificador de las explicaciones morales 170
Las explicaciones morales de algunas observaciones pueden ofrecer la ventaja de unificar fenómenos aparentemente dispares. Pero las

explicaciones de los mismos fenómenos que compiten entre sí y que invocan diferentes propiedades morales o plantean propiedades no morales unificadas pueden lograr la misma ventaja.

4.5 El argumento de la disimilitud radical 172

El argumento más simple contra el reduccionismo es que las propiedades morales parecen, a primera vista, radicalmente diferentes de las propiedades naturales.

4.6 Explicación de las creencias morales 175

Las explicaciones reduccionistas de cómo pueden justificarse las creencias morales no se aplican a las creencias reales de casi nadie.

PARTE II: EL INTUICIONISMO ÉTICO 183

5 EL CONOCIMIENTO MORAL 184

5.1 El principio del conservadurismo fenomenal 184

Es razonable suponer que las cosas son como parecen, en ausencia de motivos para dudar de ello. El juicio en general presupone este principio.

5.2 Intuiciones éticas 188

Las intuiciones se definen como apariencias intelectuales iniciales. Las intuiciones éticas son intuiciones con contenidos valorativos. No son meras creencias o productos de creencias. No todas las intuiciones son igualmente creíbles.

5.3 Malentendidos del intuicionismo 193

Los intuicionistas no consideran que las intuiciones sean infalibles o imprescriptibles, ni sostienen que todo el conocimiento moral sea intuitivo.

5.4 Objeciones epistemológicas comunes 196

Respondo a las objeciones de que (1) necesitamos un argumento de que las intuiciones son generalmente fiables, (2) no tenemos forma de

comprobar las intuiciones, (3) la intuición puede utilizarse para justificar cualquier afirmación, y (4) la intuición es "rara".

5.5 La inverosimilitud del nihilismo: un argumento mooreano 208
Las premisas que utilizan los nihilistas para defender su postura son mucho menos plausibles que muchas de las afirmaciones morales que esas premisas supuestamente refutan. Deberíamos rechazar las premisas de los nihilistas antes que rechazar todas las afirmaciones evaluativas.

5.6 El realismo directo y la inversión subjetiva 212
Los realistas directos sobre la percepción sostienen que las experiencias perceptivas hacen que las creencias sobre el mundo físico estén justificadas, no por constituir signos de estados de cosas externos, sino por constituir una aparente conciencia directa de los fenómenos externos. Los intuicionistas deberían adoptar un punto de vista realista directo similar sobre la ética.

**5.7 El aislamiento del realismo moral* 219
El conocimiento a priori se explica por nuestra capacidad de captar objetos abstractos, o universales. Una comprensión adecuada de uno o más universales conduce a creencias fiables sobre las propiedades de esos universales y las relaciones entre ellos, aunque los universales no sean la causa de nuestras creencias sobre ellos.

6 DESACUERDO Y ERROR 234

6.1 La prevalencia del desacuerdo moral 235
Hay muchos desacuerdos sobre el valor, tanto entre sociedades como entre individuos dentro de una sociedad.

6.2 El veto del idiota 238
El "veto del idiota" es la idea de que una afirmación queda descalificada como objetiva por el mero hecho de que algunas personas no estén de acuerdo con ella.

6.3 ¿Pueden los intuicionistas explicar el desacuerdo? 241
Algunos objetan que los intuicionistas no pueden explicar por qué hay muchos desacuerdos sobre el valor.

6.3.1 La caricatura 241
El argumento del desacuerdo se ha apoyado a veces en una caricatura absurda según la cual los intuicionistas sostienen que la intuición resuelve inmediata e infaliblemente todas las cuestiones morales.

6.3.2 La prevalencia del desacuerdo no moral 242
La prevalencia del desacuerdo moral no es sorprendente por el hecho de que hay muchos ejemplos de desacuerdos no morales con características similares a las de los desacuerdos morales.

6.3.3 Una colección de errores 247
Hay muchas causas de error, tanto en asuntos morales como no morales, como la parcialidad, la confusión, las falacias, los juicios precipitados, etc.

6.3.4 El desacuerdo es predecible 251
Los errores son especialmente comunes en los ámbitos en los que existen fuertes y frecuentes prejuicios, en los que las personas se atienen a sus culturas y en los que las personas se atienen a la religión, y en todos los ámbitos de la filosofía. Esto hace que el error y el desacuerdo en la ética no sean sorprendentes.

6.4 ¿Pueden los intuicionistas resolver los desacuerdos? 254
Algunos objetan que los intuicionistas no tienen forma de resolver los desacuerdos éticos.

6.4.1 Desacuerdos hipotéticos 255
Los intuicionistas no pueden ofrecer ninguna forma razonada de resolver un desacuerdo si se estipula que cada parte se niega a conceder ninguna premisa a su oponente. Pero esto no muestra nada interesante.

6.4.2 Cuestiones morales controvertidas 255
Hay algunas controversias morales reales que los intuicionistas no pueden resolver. Pero esto es cierto para todas las teorías metaéticas y no es una prueba contra el intuicionismo.

6.4.3 Controversias morales fundamentales 256
Puede haber desacuerdos sobre los principios morales fundamentales que los intuicionistas no puedan resolver. Esta preocupación es generalmente exagerada y tampoco aporta pruebas contra el intuicionismo.

6.4.4 Controversias en general 258
Algunos objetan que los intuicionistas no pueden resolver ningún desacuerdo que implique intuiciones conflictivas. Pero aunque los intuicionistas no pueden resolver todas las disputas, ni pueden obligar a la gente a ser racional, pueden ofrecer formas racionales de intentar resolver muchos de esos desacuerdos.

6.5 El problema de la autorrefutación 261
Los defensores del argumento del desacuerdo se enfrentan al problema de que muchas personas están en desacuerdo con cada teoría metaética alternativa y con el propio argumento del desacuerdo.

6.6 El desacuerdo como argumento a favor del realismo moral 264
El desacuerdo plantea un reto mayor a los antirrealistas que a los realistas.

6.6.1 ¿Pueden los antirrealistas explicar el desacuerdo? 264
Los antirrealistas tienen más problemas para explicar el desacuerdo ético que los realistas morales.

6.6.2 ¿Pueden los antirrealistas resolver los desacuerdos? 266
Los antirrealistas no pueden ofrecer ninguna forma racional de resolver los desacuerdos morales fundamentales. La práctica actual de la filosofía moral parece presuponer el intuicionismo.

7 RAZONES PRÁCTICAS 278

7.1 El argumento humeano contra el realismo 278
Las razones motivadoras se distinguen de las razones normativas. Los humeanos creen que las razones para la acción dependen de los deseos, que las actitudes morales proporcionan intrínsecamente razones para la acción, y que ninguna mera creencia sobre un hecho objetivo es suficiente para tener un deseo. Concluyen que las actitudes morales no son creencias sobre hechos objetivos.

7.2 La conexión entre las razones motivadoras y las normativas 282
El principio del "deber implica poder" y el principio de caridad interpretativa pueden utilizarse para establecer un estrecho vínculo entre las razones normativas y las motivadoras.

7.3 Una concepción racionalista de la motivación 287
Las personas están motivadas por apetitos, emociones, consideraciones prudenciales y razones imparciales. Las dos últimas no son deseos en el sentido ordinario. Las razones morales son una especie de razones imparciales.

7.4 ¿Por qué creer en la concepción humeana? 294

7.4.1 El atractivo intuitivo de la concepción humeana 294
El atractivo intuitivo de la concepción humeana de las razones depende de que sólo se tengan en mente creencias típicas no evaluativas. El argumento central de Hume para su posición falla.

**7.4.2 El argumento de Smith* 298
El argumento de Smith a favor de la concepción humeana, basado en la noción de "dirección de ajuste", se apoya en un equívoco.

7.5 ¿Extendiendo la concepción humeana? 302
Los humeanos modernos rechazan la posición extrema de Hume de que ninguna acción puede ser racional o irracional. Proponen varias

limitaciones a la acción racional.

7.5.1 La restricción de la previsión 305
La restricción de la previsión sostiene que uno está racionalmente obligado a dar cierta importancia a sus deseos futuros. Esto no puede sostenerse en una concepción ampliamente humeana de las razones normativas.

7.5.2 La restricción de la imaginación 308
La restricción de la imaginación sostiene que uno debe dar racionalmente cierto peso a los deseos que tendría si imaginara vívida y correctamente ciertos estados de cosas. La concepción humeana no puede apoyar esto.

7.5.3 Las restricciones de consistencia y coherencia 310
Las restricciones de consistencia y coherencia sostienen que uno debe modificar racionalmente un conjunto inconsistente o incoherente de deseos para hacerlo consistente y coherente. La concepción humeana no puede apoyar esto.

7.5.4 La restricción de la deliberación 314
La restricción de la deliberación sostiene que uno debe deliberar racionalmente sobre lo que debe hacer y atenerse a los resultados de esa deliberación. La concepción humeana no puede apoyar esto. La concepción humeana no permite criticar racionalmente ninguna elección.

7.5.5 Coda: la necesidad de hechos evaluativos 318
Los problemas a los que se enfrenta la concepción humeana pueden evitarse reconociendo hechos objetivos y evaluativos. Esto nos lleva al intuiciónismo. El intento de Scanlon de evitar la necesidad de hechos evaluativos no naturales se enfrenta a problemas de circularidad y regresión infinita.

7.6 La autoridad de la moral 322

7.6.1 La autoridad de la moralidad: una visión racionalista 322
Los juicios morales racionales ponderan el interés propio entre otros factores, al igual que los juicios prudenciales racionales ponderan la inclinación presente entre otros factores. En consecuencia, la moralidad prevalece sobre la prudencia, y la prudencia prevalece sobre el deseo actual.

7.6.2 La arbitrariedad de la moralidad en el punto de vista humeano 326
Para los humeanos, los imperativos morales provienen del subconjunto de nuestros deseos, seleccionados arbitrariamente, que no tienen más fuerza para dar razón que cualquier otro deseo.

7.6.3 El problema de la debilidad de la voluntad 328
Los racionalistas pueden dar cabida a la debilidad de la voluntad y al libre voluntad, mientras que los humeanos no pueden.

7.6.4 Cómo el antirrealismo desvirtúa la moral 334
El antirrealismoral desvirtúa las creencias morales y la motivación moral. Nos enseña que no tenemos motivaciones imparciales para la acción y que no hay razón para actuar moralmente cuando esto no satisface nuestros propios deseos.

7.7 ¿Por qué ser moral? 340
En una interpretación, "¿Por qué ser moral?" es una pregunta sin sentido. La naturaleza de la evaluación moral hace que la acción moral sea inherentemente racional. La pregunta "¿Por qué ser racional?" no tiene sentido.

8 OTRAS OBJECIONES 350

8.1 El argumento de la rareza 350
Mackie objeta los hechos morales por su "rareza". Pero en ninguna interpretación obvia es cierto tanto que los hechos morales sean raros como que esta rareza constituya una prueba contra su existencia.

*8.2 *Problemas con la superveniencia* 354
Las propiedades evaluativas de una cosa parecen depender de sus propiedades no evaluativas. Algunos encuentran esta relación problemática.

*8.2.1 *La objeción de Mackie* 354
Mackie considera que la relación es extraña. Esto no tiene más fuerza que su argumento general de la rareza.

*8.2.2 *Objeciones a la noción de rectitud prima facie* 356
Strawsoll considera incoherente la noción de que un tipo de acción tenga necesariamente una tendencia a ser correcta. Los ejemplos de las fuerzas en física y de las buenas jugadas en ajedrez demuestran la coherencia de la noción.

*8.2.3 *La objeción de Blackburn* 360
Blackburn encuentra desconcertante que el principio general de la superveniencia de lo ético sea analítico pero que ninguna teoría ética en particular sea analítica. Muestro que los realistas son más capaces de explicar este supuesto dato que el propio Blackburn.

*8.2.4 *La objeción de Jackson* 362
Jackson plantea un rompecabezas sobre por qué "bueno" se referiría a una propiedad evaluativa no natural, incluso si tal propiedad existiera, en lugar de a una disyunción de propiedades naturales. Pero Jackson plantea la cuestión contra el intuicionismo.

*8.2.5 *Una analogía: la superveniencia de las propiedades lógicas* 364
La superveniencia de las propiedades valorativas sobre las propiedades no valorativas es análoga a la superveniencia de las propiedades lógicas sobre las propiedades senlógicas. Los argumentos no demostrativos proporcionan una analogía especialmente estrecha.

8.3 ¿Cómo podemos entender el "bien"? 366

La inobservabilidad de la bondad, junto con los desacuerdos sobre lo que es bueno, plantean un problema sobre cómo los niños podrían aprender primero el significado de "bueno".

8.4 ¿Es el intuicionismo demasiado subjetivo? 367
Algunos han interpretado erróneamente el intuicionismo como una doctrina subjetivista. Pero el conocimiento moral no es más subjetivo, según el intuicionismo, que cualquier otro tipo de conocimiento.

8.5 ¿Mis argumentos demuestran demasiado? 369
Algunos argumentos contra el antirrealismo moral parecen aplicarse también al antirrealismo sobre la diversión, la frialdad y similares. Pero reflexionando, un relato subjetivista de estas últimas propiedades es más convincente que un relato subjetivista del valor.

8.6 Evolución y ética 373

8.6.1 La objeción evolutiva al realismo 373
Las teorías sociobiológicas sobre el origen de las actitudes morales ponen en duda la creencia de que estas actitudes siguen a los hechos objetivos.

8.6.2 La carga de los realistas 375
El conocimiento moral puede ser un subproducto de la razón. Los realistas morales no necesitan argumentar que la evolución predice la existencia del conocimiento moral, sino sólo que no predice la ausencia de conocimiento moral.

8.6.3 ¿En qué medida son buenas las explicaciones evolutivas de la ética? 377
Los planteamientos evolutivos existentes sobre la ética son poco impresionantes, tanto porque no dan cuenta de los detalles de nuestras intuiciones y creencias morales, como porque la flexibilidad del paradigma hace que las hipótesis evolutivas sean difíciles de falsar.

8.6.4 Una explicación evolutiva de la percepción moral 379

La percepción moral precisa podría tener un valor de supervivencia en virtud de que facilita la cooperación pacífica.

8.7 ¿Es la intuición demasiado revisionista? 381

Los intuicionistas pueden tener que rechazar gran parte de la moral convencional, lo que lleva a sospechar que han malinterpretado los conceptos morales ordinarios. Pero puede ser que la mayoría de la gente esté habitualmente confundida sobre la ética. Además, el intuicionismo puede ser interesante incluso si implica una reinterpretación del discurso moral.

9 CONCLUSIÓN 393

9.1 Los fracasos de las teorías alternativas de la metaética 394

9.1.1 El no cognitivismo 394

Resumo los argumentos lingüísticos e introspectivos contra el no cognitivismo.

9.1.2 Subjetivismo 397

Repaso de los cinco principales problemas a los que se enfrentan los subjetivistas.

9.1.3 Naturalismo 399

Repaso el argumento de la pregunta abierta contra el reduccionismo analítico, seguido de dos argumentos principales contra el reduccionismo sintético.

9.1.4 Nihilismo 402

Repaso de la inverosimilitud del nihilismo.

9.2 El punto de vista intuicionista 403

9.2.1 Intuición y conocimiento moral 404

Repaso mi explicación del conocimiento moral, incluido el principio del conservadurismo fenomenal.

9.2.2 La rnotivación moral 406
Repaso de la visión racionalista de las razones para la acción, así como de las objeciones al relato humeano en competencia.

9.3 Algunas objeciones al intuicionismo 411
Repaso de algunas de las objeciones más comunes al intuicionismo y mis respuestas a ellas.

9.3.1 "No se puede demostrar que la intuición sea fiable" 411

9.3.2 "Los intuicionistas no pueden ex.plicar el desacuerdo" 412

9.3.3 "Los intuicionistas no pueden resolver los desacuerdos" 415

9.3.4 "El intuicionismo es raro" 416

9.4 La revuelta contra los valores: cómo el intuicionismo perdió el favor 418
Explico por qué el intuicionismo es impopular.

9.4.1 ¿Fue el intuicionismo refutado racionalmente? 418
El intuicionismo no recibió un examen justo a finales del siglo XX. Más bien fue rechazado porque no encajaba con el espíritu de la época.

9.4.2 El cinismo 420
Los intelectuales modernos han tratado de derribar los valores de su propia sociedad. El intuicionismo no encaja con el cinismo moderno.

9.4.3 Lo políticamente correcto 422
El intuicionismo es políticamente incorrecto. Sugiere que algunas personas son moralmente malas y que algunas culturas tienen creencias muy equivocadas.

9.4.4 El cientificismo 234
El intuicionismo no encaja con el culto moderno a la ciencia. Las propiedades morales no son estudiadas por la ciencia, y la "intuición"

no se parece al método científico.

9.4.5 Cómo el intuicionismo se vuelve "inverosímil" 426

Los factores anteriores sesgan el juicio de los intelectuales cuando evalúan las teorías metaéticas. Las actitudes de los intelectuales se filtran al resto de la sociedad.

9.5 La importancia del intuicionismo 429

El intuicionismo es un importante contrapeso a las teorías antirrealistas que socavan nuestra motivación moral y nuestro sentido del valor y el significado de la vida.

9.6 Cómo me convertí en un intuicionista espeluznante y anticientífico 433

Mi camino hacia el intuicionismo pasó por darme cuenta, primero, de que la mayor parte del conocimiento es indemostrable, a diferencia del conocimiento científico y matemático; y, segundo, de que la dependencia de las propiedades morales de las no morales no implica la existencia de un algoritmo para computar veredictos morales a partir de hechos no morales.

Prefacio

Todos emitimos juicios de valor, pero casi ninguno entiende lo que hace cuando lo hace. A través de conversaciones y debates que he mantenido durante varios años, he llegado a la conclusión de que casi todos los intelectuales de nuestra sociedad piensan que la moral es de alguna manera irreal. He llegado a esperar, cada vez que surge el tema de la naturaleza de los valores, que me digan alegremente que la moralidad es todo una cuestión de emociones o convenciones, que todo es una ilusión creada por nuestros genes, o que es un mito patrocinado por la religión... Este parece ser el punto de vista sofisticado y "científico". Hace poco encuesté a una clase de cuarenta estudiantes universitarios sobre el tema. Tras explicar los términos "subjetivo" y "objetivo", pregunté cuántos de ellos creían que "la moral es subjetiva". Todos los alumnos levantaron la mano, excepto dos, que éramos mi ayudante y yo. Esto es aún más notable por el hecho de que normalmente es casi imposible lograr un acuerdo universal, en una clase de filosofía, sobre cualquier cosa. Los profesores de filosofía, cuyo trabajo es estudiar cosas como la naturaleza de los valores, están menos unidos. El objetivismo sigue siendo una posición minoritaria respetable en el campo. Sin embargo, la mayoría de los expertos parecen estar de acuerdo en que la moral es, en cierto sentido, irreal.

Nada de esto parece impedir que nadie -ya sean estudiantes, profesores u otros intelectuales- emita juicios morales, discuta sobre cuáles son los puntos de vista morales correctos o intente que los demás

obedezcan los principios morales correctos. Incluso aquellos que declaran que la moral es una ilusión, a menudo proceden a argumentar sobre lo erróneo de la guerra de Irak, o de la clonación humana, o al menos de que su novio les engañe. Y parecen esperar que sus argumentos se tomen como razones para que otras personas actúen de determinada manera. Esto me parece extraño. Si yo pensara que el conejo gigante que está en la esquina de la habitación es una alucinación, no creo que me pusiera a hablar en público sobre cuál es su comida favorita, ni planearía mis acciones en torno a su horario, ni esperaría que los demás modificaran su comportamiento a la luz de mis afirmaciones sobre él. Si la moral es una ilusión, tampoco está claro por qué alguien debería preocuparse por sus alucinantes dictados. Y los que consideran la moral como una cuestión de convenciones o de emociones no parecen tratarla en la práctica como tal. No discuten sobre lo que es moral como se esperaría que se discutiera sobre cuáles son las convenciones sociales o qué emociones siente la gente. Parece que tratan sus afirmaciones morales como si tuvieran algún tipo de fuerza mayor que las afirmaciones sobre las convenciones o las emociones. Si el aborto no se ajusta a las convenciones sociales estadounidenses, o si estimula emociones negativas en ciertos observadores, ¿por qué se supone que eso va a convencer a una mujer embarazada que no quiere tener un bebé de que lo lleve a término de todos modos?

Tal vez mis preguntas sean ingenuas, y quizá los antirrealistas morales tengan algún tipo de respuesta sofisticada para ellas. Yo sólo informo de lo que me parece a primera vista. A primera vista, uno pensaría que el descubrimiento de la filosofía moderna -si es que se trata de eso- de que la moral es subjetiva, ilusoria o no objetiva tendría un profundo impacto en cómo pensamos y hablamos sobre cuestiones morales; sin embargo, los que abrazan el supuesto descubrimiento en un instante parecen olvidarlo al siguiente, sin dedicar casi ninguna reflexión

a las implicaciones que podría tener para la práctica de la argumentación moral, la exhortación, etc. Se sugiere una explicación sencilla: quizás la mayoría de las declaraciones de antirrealismo son fundamentalmente insinceras. En el contexto de una discusión filosófica abstracta, decimos que la moral es irreal, e incluso podemos decirnos a nosotros mismos que lo creemos. Pero lo que realmente creemos se revela más por la forma en que hablamos de la moral en situaciones concretas y por la forma en que ordenamos nuestras vidas de acuerdo con los principios morales que por lo que decimos en la sala de filosofía.

Pero no es tan sencillo. A veces, el antirrealismo moral afecta a la forma de hablar de las cuestiones morales. La gente argumentará que el gobierno no debería "legislar la moral" porque la moral es subjetiva. O que no deberíamos intentar evitar la circuncisión femenina porque la moral es culturalmente relativa. O que deberíamos abstenernos de juzgar a los demás, o que un profesor no debería presumir de enseñar principios morales, porque no se conocen verdades morales objetivas. Por supuesto, la mayoría de los filósofos profesionales se avergonzarían de escuchar tales argumentos. Si la moral es subjetiva, no se deduce que el gobierno no deba legislarla; lo que se deduce es que el gobierno debe legislar la moral si lo hace de acuerdo con las preferencias subjetivas de los legisladores. Si la moralidad es culturalmente relativa, no se deduce que no debamos interferir en las costumbres de otras culturas; lo que se deduce es que debemos interferir en otras culturas si hacerlo concuerda con nuestras costumbres. Si no se conocen las verdades morales objetivas, no se deduce que debamos abstenernos de juzgar a los demás o de enseñar principios morales; lo que se deduce es que no sabemos si es objetivamente cierto que debamos juzgar a los demás o enseñar principios morales. Los tres argumentos mencionados al principio de este párrafo parecen proceder eximiendo inconscientemente algún principio moral del antirrealismo general que presuponen: el principio de que uno

no debe hacer leyes basadas en meras preferencias subjetivas; que uno no debe imponer convenciones parroquiales a otras sociedades; que uno no debe hacer juicios que no sabe que son objetivamente correctos; que uno no debe enseñar cosas que no sabe que son objetivamente correctas.

Por lo tanto, estoy a favor de una explicación diferente de las actitudes de nuestra cultura hacia la moralidad: Sugiero que son incoherentes; de hecho, lo son de forma flagrante. Cualquiera que sea el pensamiento de la mayoría de los individuos sobre la naturaleza del valor no resistiría un minuto de escrutinio. Pensamos que los valores son subjetivos, pero que la guerra de Irak fue objetivamente errónea; pensamos que la moral es una ilusión, pero que todos deberíamos actuar moralmente; pensamos que, como no hay valores objetivos, es objetivamente erróneo imponer nuestros valores a los demás.

La pregunta es: ¿cuáles de nuestras creencias contradictorias son falsas y cuáles, si es que hay alguna, son correctas?

La mayoría de los que niegan la existencia de valores objetivos reconocerán que, al menos a primera vista, parece natural suponer que hay valores objetivos. Casi todas las sociedades a lo largo de la historia han dado por sentada la objetividad de los valores. Y como he sugerido, incluso los miembros de nuestra propia sociedad cínica parecen asumir la objetividad de los valores en su pensamiento ordinario sobre cuestiones morales concretas. Si realmente no hay valores objetivos, entonces éste debe ser el descubrimiento más significativo de la filosofía moderna, y quizás la primera vez que la disciplina de la filosofía ha conseguido convencer a un gran número de personas para que abracen una revisión masiva del sentido común. Si, por el contrario, hay valores objetivos, entonces la opinión generalizada de lo contrario debe ser uno de los mayores errores de la filosofía moderna, y de la cultura intelectual moderna en general.

Esto último es lo que yo creo. He escrito este libro para defender un relato completamente objetivista y racionalista de la naturaleza de la moral y del conocimiento moral. El punto de vista que defiendo se conoce, de forma algo engañosa dadas las connotaciones del término "intuición" en la cultura popular, como intuición ética. Sostiene que hay hechos evaluativos objetivos -hechos como que está mal causar sufrimiento gratuito a los demás- por encima de los hechos naturales no evaluativos; que tenemos una especie de visión intelectual de algunos de estos hechos evaluativos; y que nos proporcionan razones para comportarnos de determinadas maneras, independientemente de lo que deseemos. Esta postura se considera en general ingenua e indefendible. Yo creo, por el contrario, que las objeciones comunes a la misma son mucho más débiles de lo que se ha considerado y no podrían haber movido a ningún intuicionista razonablemente reflexivo a abandonar su posición.

La primera parte del libro, que sigue al capítulo introductorio, es negativa: trata de refutar tres teorías alternativas sobre el valor. La segunda parte explica y defiende mis propios puntos de vista sobre el valor: el capítulo 5 explica cómo conocemos las verdades morales; el capítulo 6 trata el problema del desacuerdo y el error moral; el capítulo 7 explica cómo los valores proporcionan razones para la acción; y el capítulo 8 responde a numerosas objeciones. Finalmente, el capítulo 9 ofrece una revisión de los principales argumentos del libro, junto con algunas especulaciones sobre por qué las conclusiones que defiendo son impopulares y por qué las cuestiones son importantes.

¿Quién debería leer este libro? He tratado de escribir un libro que pueda ser leído con provecho por otros profesores, pero no he tratado de escribir uno que sólo pueda ser leído por profesores. La naturaleza de la moral y los valores es un asunto de todos. Los problemas del relativismo

moral y del escepticismo, si es que son problemas, afectan a los estudiantes y al lector cualquiera tanto como a los filósofos profesionales. Así que he dirigido mi trabajo tanto a los filósofos profesionales como a los aficionados. Se trata de una empresa difícil, y sin duda habrá opiniones divergentes sobre el éxito de mi trabajo. He tratado de hacer avanzar el estado del campo, pero también he explicado argumentos clásicos que los estudiantes nuevos en el campo deberían escuchar. Algunos de mis colegas pueden aburrirse ocasionalmente por la repetición de viejos argumentos, mientras que algunos lectores pueden sentirse confundidos por los puntos técnicos. Para minimizar esta última dificultad, he marcado con asteriscos ("*") las secciones más técnicas del libro, incluidas algunas secciones que responden a puntos de vista expuestos por personas concretas del mundo académico contemporáneo. El no especialista puede saltarse estas secciones sin perder el hilo argumental.

Este libro se ha beneficiado de los comentarios de varios amigos y colegas, como Elinor Mason, Doug Husak, Ari Armstrong, Bryan Caplan, Robin Hanson, Tyler Cowen, Ananda Gupta y dos revisores anónimos de Oxford University Press. Stuart Rachels y Richard Fumerton merecen un reconocimiento especial por sus extensos e inestimables comentarios sobre el manuscrito. Estoy agradecido a todas estas personas, sin las cuales el libro sería mucho menos satisfactorio de lo que es.

M.H

1

Introducción

1.1 El campo de la metaética

El campo de la ética se ocupa de las preguntas evaluativas: son preguntas sobre lo que es bueno o malo, o sobre lo que debe o no debe hacerse. Una pregunta evaluativa requiere un enunciado evaluativo como respuesta: se trata de un enunciado que intrínsecamente hace una evaluación positiva o negativa de algo.[1] Los siguientes son ejemplos de enunciados evaluativos: "La felicidad es buena"; "Pol Pot era malo"; "La honestidad es una virtud". La última es evaluativa porque forma parte del significado de "virtud" que las virtudes sean buenas.

La metaética, por su parte, es la rama de la filosofía que se ocupa de las cuestiones relativas a la naturaleza de los valores y las afirmaciones evaluativas. Por lo general, estas preguntas no son en sí mismas evaluativas, es decir, no exigen afirmaciones evaluativas como respuesta. He aquí algunos ejemplos.

En primer lugar, hay cuestiones semánticas: Cuando digo: "Está mal comerse a tu perro", ¿qué significa eso?[2] ¿Se puede definir "mal"? ¿Se refiere a una propiedad que supuestamente tiene el comer perros? ¿Se refiere "bueno" a una propiedad que tienen la vida, la felicidad, etc, y

puede definirse?

En segundo lugar, hay cuestiones epistemológicas: ¿Cómo sabemos, si es que lo sabemos, que está mal comerse al perro? ¿Podemos justificar nuestras creencias evaluativas con argumentos? ¿Necesitamos hacerlo? ¿Tenemos alguna forma de "percibir" los valores morales?

En tercer lugar, hay al menos una cuestión metafísica importante: ¿Existen valores objetivos? O bien: ¿Son algunas afirmaciones evaluativas objetivamente verdaderas?

En cuarto lugar, hay preguntas sobre la psicología moral y la acción racional: ¿Qué nos motiva a actuar (normalmente) de la manera que consideramos moral? ¿Qué buenas razones, si las hay, tenemos para ser morales?

Todas estas son cuestiones metaéticas, no éticas. Obsérvese que, aunque todas ellas tienen que ver con el valor de un modo u otro, generalmente no son preguntas evaluativas. Cuando pregunto: "¿Existen valores objetivos?", la respuesta es: "Existen valores objetivos" o "No existen valores objetivos". Ninguna de estas respuestas es en sí misma una afirmación evaluativa; ninguna expresa un juicio positivo o negativo sobre algo. Cuando pregunto: "¿Tenemos alguna forma de "percibir" los valores morales?", la respuesta es que tenemos esa capacidad o que no la tenemos; ninguna de estas respuestas es en sí misma un juicio de valor.

Este libro trata de ese tipo de preguntas. Mi objetivo es mostrar que existen valores objetivos, explicar cómo los conocemos y cómo nos dan razones para actuar, y refutar a los principales competidores de mis propios puntos de vista en metaética.

1.2 ¿Qué es la objetividad?

La cuestión metaética más discutida es la de si el valor es "objetivo". ¿Qué significa que un fenómeno sea "objetivo"?

Cuando una persona es consciente de algo, la llamamos "sujeto" de la conciencia. La cosa de la que es consciente la llamamos "objeto" de la conciencia.[3] Que una característica de algo sea "objetiva" significa, a grandes rasgos, que está en el objeto. Un rasgo "subjetivo" es el que está en el sujeto.[4] Pero, ¿qué significa estar "en el sujeto"? Propongo esta definición:

> La *F*-sidad es subjetiva = Que algo sea *F* constitutivamente depende, al menos en parte, de la actitud o respuesta psicológica que los observadores tienen o tendrían hacia esa cosa.

Defino una característica "objetiva" como aquella que no es subjetiva.[5] Consideremos algunos ejemplos:

La gracia es subjetiva, porque el hecho de que un chiste sea gracioso depende de que a la gente le haga gracia. Si nadie lo encuentra divertido, entonces no lo es. Este es un punto sobre el significado de "gracioso": "gracioso" significa algo así como "tiene tendencia a divertir al público".

Del mismo modo, la sensualidad es subjetiva, porque lo que es sexy depende de lo que la gente se sienta atraída. Esto, de nuevo, es una cuestión del significado de "sexy".

En cambio, la cuadratura es objetiva: que un objeto sea cuadrado no tiene nada que ver con las reacciones de los observadores. Para ser cuadrada, una figura sólo tiene que tener cuatro lados iguales y ángulos iguales. Nadie tiene que sentir nada al respecto, ni pensar en ello, ni siquiera verlo.

¿Y los rasgos psicológicos, como la felicidad, son subjetivos? No en el sentido que pretendo. Que una persona sea feliz depende de las

actitudes de alguien, es decir, de la propia persona, pero no depende de las actitudes de los observadores hacia ella. Que uno sea feliz no depende de lo que sienta quien le observa.

Otra noción que hay que aclarar es la de dependencia constitutiva. Consideremos este ejemplo. Existe la creencia generalizada de que, en el caso de algunas enfermedades, las posibilidades de recuperación de un paciente se ven afectadas por su actitud: si un paciente adopta una actitud positiva, esto tiene un efecto beneficioso en su cuerpo, lo que hace que tenga más probabilidades de recuperarse. Imaginemos una enfermedad que está excepcionalmente sujeta al poder del pensamiento positivo: quien cree que se recuperará de la enfermedad lo hace, y quien cree que no se recuperará no se recupera. ¿Debemos decir entonces que la recuperación o no recuperación de la enfermedad de un individuo determinado es "subjetiva", porque el hecho de que se recupere depende de que crea que lo hará?

No; ese es un caso de dependencia causal (la actitud positiva del paciente provoca su recuperación), no de dependencia constitutiva. Una propiedad subjetiva es aquella que está constituida, al menos en parte, por su tendencia a provocar una determinada reacción en los observadores. En otras palabras, si la *F*-sidad es "subjetiva", entonces parte de lo que significa que una cosa sea *F* es que los observadores tengan o estén dispuestos a tener algún tipo particular de reacción ante ella. En el ejemplo anterior, resulta que hay una conexión causal entre el pensamiento positivo y la recuperación; no es que parte de lo que constituye la recuperación de la enfermedad sea la actitud positiva de uno (lo que constituiría la recuperación es el cese de cualquiera de los síntomas de la enfermedad). Por el contrario, la capacidad de divertir a la gente es (todo o parte de) lo que constituye ser gracioso. Por eso la diversión es una propiedad subjetiva en mi sentido, mientras que la

"propiedad" de recuperarse de nuestra hipotética enfermedad no lo es.

Hay casos controvertidos. Algunos filósofos piensan que los colores son subjetivos. Algunos piensan, por ejemplo, que para que una cosa sea azul basta con que esté dispuesta a provocar una determinada sensación (la "sensación de azul") en los humanos normales que lo ven en condiciones normales. Para nuestro propósito, no importa si esto es correcto o no. Lo que importa es entender qué se entiende por una teoría subjetivista. Se trata de una teoría subjetivista del color, ya que hace que el color de un objeto físico dependa constitutivamente de las respuestas psicológicas de los observadores a él (contando las sensaciones como "respuestas psicológicas").

1.3 Cinco teorías metaéticas

¿Es objetiva la bondad? Los realistas morales dicen que sí; los antirrealistas morales dicen que no. Hay tres formas de antirrealismo:

1. El *subjetivismo* (que incluye el relativismo) sostiene que las propiedades morales son subjetivas en el sentido que acabamos de definir: para que una cosa sea buena es necesario que algún individuo o grupo adopte (esté dispuesto a) alguna actitud hacia ella. La forma más simple de subjetivismo afirma que "*x* es bueno" significa "apruebo *x*". Obsérvese que, dado que en este análisis, "bueno" se aplica a cualquier cosa que el hablante apruebe, una persona podría decir con verdad: "La poligamia es buena", mientras que otra dice con verdad: "La poligamia no es buena". Sus afirmaciones no entran en conflicto lógico, como tampoco lo harían si el primero dijera: "Me gustan los pepinillos" y el otro dijera: "No me gustan los pepinillos".

Otras formas de subjetivismo sustituyen la aprobación por otras actitudes y al orador por otras personas o grupos. Así, la teoría de que ser bueno es ser aprobado por la sociedad es una forma de subjetivismo, al igual que la teoría de que ser obligatorio es ser mandado por Dios.

2. El *no cognitivismo* sostiene que los predicados evaluativos ni siquiera pretenden referirse a ningún tipo de propiedad,[6] ni los enunciados evaluativos afirman proposiciones. Los enunciados evaluativos no hacen afirmaciones sobre cómo es el mundo, ni siquiera sobre la parte del mundo que incluye a los observadores humanos. ¿Qué hacen en cambio? Algunos dicen que sólo expresan las emociones de los hablantes. Así, "Stalin era malvado" es como "¡Buu Stalin! Obsérvese que "¡Bu!" no es una afirmación (no dice que el mundo sea de una manera determinada), ni siquiera una afirmación sobre los sentimientos del hablante, aunque puede utilizarse para expresar los propios sentimientos. Por eso "¡Buu!" no es ni verdadero ni falso. Del mismo modo, los no cognitivistas han negado tradicionalmente que los enunciados evaluativos puedan ser verdaderos o falsos.[7]

   Otra variante del no cognitivismo sostiene que los enunciados evaluativos se parecen más a imperativos. "Robar está mal" es como "¡No robes!". Obsérvese que, de nuevo, "No robes" no es una afirmación y no es ni verdadero ni falso.

3. El *Nihilismo* (también conocida como "la teoría del error") sostiene que los enunciados evaluativos son generalmente falsos.[8] ¿Por qué? Porque los enunciados evaluativos afirman que las cosas tienen propiedades de valor objetivas, pero en realidad no existen tales propiedades. Cuando decimos "Stalin era malvado", no sólo queremos decir que no nos gusta Stalin, ni

> tampoco estamos tratando de expresar nuestros sentimientos hostiles hacia él. Pensamos que existe esta propiedad de "maldad", y que él la tenía, y de eso trata el lenguaje moral. Sin embargo, dice el nihilista, estamos equivocados: realmente no hay ninguna propiedad así. Así que nuestra afirmación es simplemente falsa. Nada es bueno ni malo.

En capítulos posteriores veremos por qué los filósofos han mantenido estas opiniones. En este momento, es importante ver por qué las tres posiciones anteriores son las únicas formas posibles de antirrealismo. O bien "bueno" pretende referirse a una propiedad -es decir, "*x* es bueno" afirma que *x* tiene una determinada propiedad- o bien no la tiene. Si la tiene, entonces o algunas cosas tienen la propiedad o nada la tiene, y o la propiedad depende de los observadores o no depende de ellos. Si decimos que "bueno" no pretende referirse a una propiedad en absoluto, entonces tenemos el no cognitivismo ético. Si decimos que "bueno" pretende referirse a una propiedad pero nada tiene esa propiedad, entonces tenemos nihilismo. Si decimos que "bueno" pretende referirse a una propiedad, algunas cosas tienen esa propiedad, pero la propiedad depende de los observadores, entonces tenemos subjetivismo. Por último, si decimos que "bueno" pretende referirse a una propiedad, algunas cosas tienen esa propiedad, y la propiedad no depende de los observadores, entonces tenemos el realismo moral. Ésas son las únicas posibilidades.

Podemos hacer lo mismo -que hay exactamente tres posiciones antirrealistas posibles- considerando la cuestión de las "verdades morales objetivas". Un enunciado es objetivamente verdadero si y sólo si: es verdadero y lo que lo hace verdadero no son ni siquiera en parte las actitudes o reacciones psicológicas de los observadores hacia las cosas de las que trata el enunciado. Así, "Los mosquitos tienen 47 dientes" es

objetivamente verdadera, porque es verdadera y lo que la hace verdadera no es ni siquiera en parte mi reacción o la de cualquier otra persona hacia las cosas (mosquitos) de las que trata la afirmación. Pero "Jon Stewart es divertido" no es objetivamente cierto: es cierto, pero lo que lo hace cierto es en parte la reacción que la gente tiene hacia la cosa (Jon Stewart) de la que trata la afirmación.

Los antirrealistas niegan la existencia de verdades morales objetivas. Si una afirmación moral no es objetivamente verdadera, entonces sólo hay tres posibilidades: o es no objetivamente (subjetivamente) verdadera, o es falsa, o no es ni verdadera ni falsa. Por tanto, tenemos de nuevo las posturas subjetivista, nihilista y no cognitivista.[9]

El hecho de que éstas sean las tres únicas formas posibles de antirrealismo es importante, ya que significa que si podemos refutarlas todas, habremos establecido el realismo moral.

El realismo moral se presenta en dos variedades principales:

1. El *Naturalismo Ético* sostiene que existen propiedades morales objetivas, pero que son reducibles. Tradicionalmente, los naturalistas pensaban que se podían definir (es decir, explicar el significado de) terminos como "bueno" utilizando términos totalmente no valorativos. Por ejemplo, tal vez el bien pueda definirse como cualquier cosa que promueva la supervivencia, la salud y la felicidad humanas.

   A finales del siglo XX, se impuso otra forma de naturalismo, una forma que no afirma que el significado de la palabra "bueno" pueda explicarse en términos no valorativos, pero que, sin embargo, sostiene que lo que es la bondad puede explicarse en términos no valorativos. Estos naturalistas modernos comparan su visión de las propiedades evaluativas con las teorías científicas sobre la naturaleza del calor, el sonido o el agua. La palabra "agua" no significa lo

mismo que "compuesto químico cuyas moléculas contienen dos átomos de hidrógeno y uno de oxígeno". Esto se puede ver en el hecho de que la gente entendía comentarios como "me gustaría un vaso de agua" mucho antes de que alguien conociera la química. Sin embargo, el agua es, de hecho, H_2O. Del mismo modo, los naturalistas modernos sostienen que será posible explicar lo que es la bondad utilizando un lenguaje no valorativo, aunque la corrección de esta explicación no se desprenderá simplemente del significado de la palabra "bueno".

Además de la cláusula de que las propiedades morales son reducibles, los naturalistas suelen proponer una segunda tesis importante: que los enunciados morales pueden justificarse empíricamente; es decir, que pueden justificarse en última instancia sobre la base de la observación.

2. El *Intuicionismo Ético* sostiene que las propiedades morales son objetivas e irreductibles. Así, "bueno" se refiere a una propiedad que tienen algunas cosas (quizás acciones, estados de cosas, etc.), independientemente de nuestras actitudes hacia esas cosas, y no se puede decir cuál es esa propiedad si no es utilizando un lenguaje evaluativo ("bueno", "deseable", "debería", "valioso" y similares).

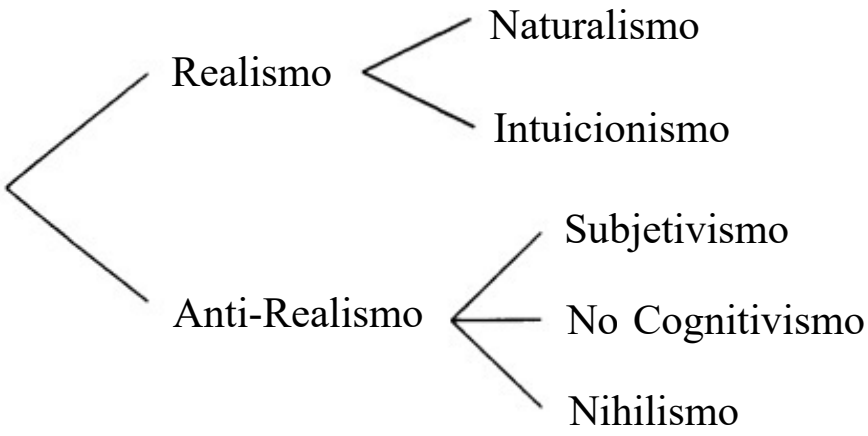
   Los intuicionistas también tienen una tesis epistemológica, de la que su doctrina recibe el nombre: que al menos algunas verdades morales se conocen de forma intuitiva. La noción de "intuición" está sujeta a interpretación, pero al menos esto es lo que generalmente se quiere decir: algunas verdades morales se conocen directamente, en lugar de sobre la base de otras verdades, pero no por los cinco sentidos (no vemos el valor moral con nuestros ojos, ni lo escuchamos con nuestros oídos, etc.).

El naturalismo y el intuicionismo difieren, pues, en dos frentes: difieren metafísicamente, sobre la cuestión de si las propiedades evaluativas son

reducibles; y difieren epistemológicamente, sobre la cuestión de si el conocimiento moral es empírico. Sería posible sostener otras formas de realismo combinando el punto de vista naturalista sobre una cuestión con el intuicionista sobre la otra. Pero ningún filósofo actual que yo conozca ha sostenido tal punto de vista, por lo que no lo discutiré en este libro. Los filósofos que han sostenido que el valor es una propiedad irreductible siempre han apelado a la intuición para dar cuenta de nuestra conciencia de esta propiedad, y los que han considerado que el valor es reducible a propiedades naturales (es decir, no valorativas) siempre han intentado dar cuenta del conocimiento moral empíricamente.

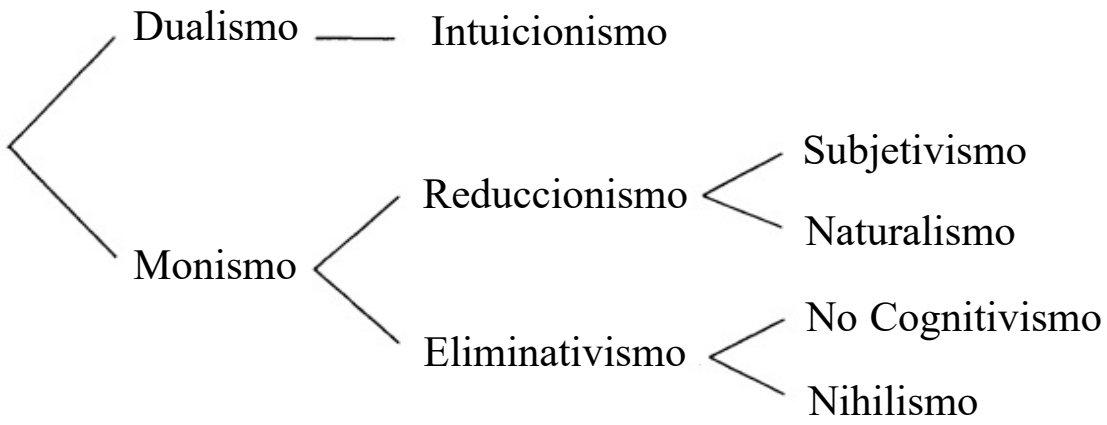
1.4 Una taxonomía alternativa de los puntos de vista metaéticos

Como sugiere la sección anterior, las teorías metaéticas se dividen tradicionalmente primero en puntos de vista realistas y antirrealistas, y luego en dos formas de realismo y tres tipos de antirrealismo:



Esta no es la forma más esclarecedora de clasificar las posiciones. Implica que la división más fundamental en la ética es entre realistas y antirrealistas sobre la cuestión de la objetividad. La disputa entre el

naturalismo y el intuicionismo se ve entonces como relativamente menor, estando los naturalistas mucho más cerca de los intuicionistas que, por ejemplo, de los subjetivistas. No es así como yo veo las cosas. Tal como yo lo veo, la división más fundamental en la metaética es entre los intuicionistas, por un lado, y todos los demás, por otro. Yo clasificaría las posiciones de la siguiente manera:



(He tomado prestada la terminología de la literatura de filosofía de la mente.) Aquí, el dualismo es la idea de que hay dos tipos fundamentalmente diferentes de hechos (o propiedades) en el mundo: hechos (propiedades) evaluativos y hechos (propiedades) no evaluativos. Sólo los intuicionistas abrazan esto.

Todos los demás son monistas: dicen que sólo hay un tipo fundamental de hechos en el mundo, y es el tipo no evaluativo; no hay hechos evaluativos por encima de los demás hechos. Esto implica que, o bien no hay hechos de valor en absoluto (eliminativismo), o bien los hechos de valor son totalmente explicables en términos de hechos no evaluativos (reduccionismo). El subjetivismo y el naturalismo se ven así como hermanos enzarzados en una disputa familiar: están de acuerdo en que podemos reducir la bondad a algo más; sólo discrepan sobre si ese algo debe incluir las reacciones psicológicas de los observadores.

¿Por qué mi taxonomía es más esclarecedora que la tradicional? Porque pone en primer plano la cuestión de lo que existe en el mundo. Los intuicionistas difieren fundamentalmente de todos los demás en su visión del mundo. Los subjetivistas, los naturalistas, los no cognitivistas y los nihilistas coinciden en su visión básica del mundo, ya que no tienen desacuerdos significativos sobre cuáles son los hechos no evaluativos, y todos están de acuerdo en que no hay más hechos por encima de éstos.[10] Están de acuerdo, por ejemplo, en las propiedades no valorativas del acto de robar, y están de acuerdo, en contra de los intuicionistas, en que no hay ninguna otra propiedad distintivamente valorativa del acto.

Entonces, ¿qué tipo de disputa tienen las cuatro teorías monistas? Creo que, aunque esto no se reconozca generalmente, sus disputas entre sí son meramente semánticas. Una vez que se ha acordado la naturaleza del mundo "ahí fuera", sólo quedan las disputas semánticas.

Es cierto que utilizo el término "disputa semántica" en un sentido algo más amplio que el habitual. No quiero decir que las cuatro teorías monistas sean variantes notacionales entre sí, o que sólo el uso diferente de las palabras por parte de sus partidarios haga que parezcan estar en desacuerdo. Lo que quiero decir es que su desacuerdo no es fundamentalmente sobre cuáles son los hechos del mundo, sino sobre la relación entre los hechos del mundo y nuestro lenguaje y pensamientos. Esto es más obvio en el caso de los no cognitivistas, que difieren de los otros monistas en la cuestión de si los enunciados evaluativos hacen afirmaciones fácticas. Se podría decir que los subjetivistas y los naturalistas difieren sobre lo que es la bondad, pero como no están en desacuerdo sobre cuáles son los hechos no evaluativos, en realidad se trata sólo de una disputa sobre cuál de las características no evaluativas de las cosas denota la palabra (o el concepto) "bueno". Incluso los nihilistas difieren de los reduccionistas por motivos puramente

semánticos: difieren sobre si los hechos no valorativos comúnmente aceptados hacen que los enunciados valorativos sean verdaderos. Los nihilistas piensan que nada es bueno porque creen que el significado de la palabra "bueno" es tal que excluye que denote cualquier propiedad natural o psicológica, mientras que los reduccionistas no están de acuerdo con esta afirmación semántica.

La clasificación tradicional parece anteponer la cuestión metafísica. La cuestión de si existen valores objetivos parece ser una cuestión metafísica, y lo sería si todas las partes estuvieran de acuerdo en el análisis del "valor". Pero teniendo en cuenta las opiniones de muchos filósofos, la cuestión es en parte metafísica y en parte semántica. Depende de (a) si los enunciados morales afirman la existencia de valores objetivos, y (b) si estos enunciados son a veces verdaderos.

Nada de esto tiene que ver con la contundencia de mis argumentos centrales en este libro; la plausibilidad de cualquiera de los cinco puntos de vista metaéticos no depende en absoluto de lo que uno considere la forma "más esclarecedora" de clasificarlos. Sin embargo, sí influye en el interés de mis afirmaciones centrales. El hecho de que el intuicionismo difiera radicalmente de todas las demás teorías de la metaética lo convierte en un punto de vista especialmente interesante, si puede defenderse.

1.5 Un intuicionismo racionalista

Mi posición es una forma de intuicionismo ético. No es la única forma posible: algunos intuicionistas estarían en desacuerdo con algunas de mis afirmaciones, aunque todos estarían de acuerdo con el núcleo de las mismas. Aquí enumero los principales principios de mi forma de intuicionismo racionalista, junto con una indicación muy breve de cómo

pretendo argumentar cada uno de ellos.

En primer lugar, hay una tesis semántica: Los predicados evaluativos como "bueno" funcionan para atribuir características objetivas a las cosas. Se trata únicamente de un punto sobre los significados de las palabras: la tesis semántica no implica que las cosas tengan propiedades de valor objetivas; sólo afirma que, cuando llamamos a una cosa "buena", estamos diciendo, correctamente o no, que tiene una propiedad de valor objetiva.

Esta tesis puede establecerse observando cómo utilizamos el lenguaje evaluativo. Por ejemplo, podemos decir: "Si el adulterio es malo, entonces Clinton debe ser castigado". Pero nunca diríamos algo como: "Si el adulterio es malo, entonces castigar a Clinton es bueno" o "Si no se comete adulterio, entonces castiga a Clinton", ya que las dos formas de no cognitivismo mencionadas anteriormente nos animarían a interpretar el enunciado. Tampoco nos parece intuitiva la inferencia "Nuestra sociedad acepta comer animales; por lo tanto, comer animales no está mal", como una forma de subjetivismo nos llevaría a esperar. Estos son sólo dos de los muchos ejemplos que se pueden dar para ilustrar que, de hecho, utilizamos el lenguaje evaluativo exactamente de la misma manera que lo haríamos si atribuyéramos propiedades objetivas a las cosas.

En segundo lugar, hay una tesis metafísica: No siempre nos equivocamos cuando hacemos juicios de valor; algunas cosas son buenas y otras malas. Por tanto, hay valores objetivos. La justificación de esto depende del siguiente punto.

En tercer lugar, hay una tesis epistemológica: Está justificado que creamos algunas afirmaciones valorativas sobre la base de intuiciones racionales. Nuestro conocimiento de las verdades morales no se deriva

totalmente de la percepción sensorial, ni las verdades morales se derivan todas de las verdades no morales. Al menos algunas verdades morales son evidentes.

En mi opinión, las intuiciones son una especie de estado mental o experiencia, distinta y normalmente anterior a la creencia, que solemos tener cuando pensamos en ciertos tipos de proposiciones, incluidas algunas proposiciones morales. Son las experiencias que reportamos cuando decimos que una cosa es "obvia" o "parece verdadera". ¿Por qué se justifica que creamos proposiciones sobre la base de tales experiencias? Sostengo que, en general, tiene sentido asumir que las cosas son como parecen, hasta que se demuestre lo contrario. Todos los razonamientos y juicios se basan en este principio, e incluso los argumentos de los antirrealistas morales se basan en intuiciones. También sostengo que muchas de las objeciones al uso de las intuiciones éticas requerirían, si se aplicaran sistemáticamente, abandonar también la percepción sensorial y renunciar a todo conocimiento del mundo.

En cuarto lugar, hay dos tesis estrechamente relacionadas sobre la naturaleza de las razones para la acción: Cuando tenemos creencias evaluativas justificadas, esas creencias pueden constituir buenas razones para la acción, razones que tenemos independientemente de nuestros deseos. Si creo justificadamente que robar está mal, entonces tengo una buena razón para no robar el caramelo de mi sobrino, independientemente de lo mucho que quiera hacerlo, de que pueda salirse con la mia, etc. Además, las creencias morales no sólo constituyen buenas razones a favor o en contra de ciertas acciones, sino que a veces también constituyen nuestros verdaderos motivos para actuar, motivos que son cualitativamente diferentes de nuestros deseos.

Estas afirmaciones entran en conflicto con la popular teoría humeana de las razones prácticas, según la cual todas las razones o motivos para

la acción dependen de los deseos. Supongamos que tengo el deseo de tomar un helado de café y pienso que ir a la tienda me ayudará a conseguirlo. Entonces tengo tanto una buena razón como una motivación para ir a la tienda. Según los humeanos, todas las razones y motivaciones prácticas implican deseos de esa manera. Así que cuando decido no robar los caramelos porque creo que robar está mal, eso también implica algún deseo mío. Quizás tengo el deseo de no romper las reglas de la sociedad, o el deseo de no ver llorar a mi sobrino, o simplemente el deseo de no ser un ladrón. En consecuencia, si tuviera otros deseos, no tendría ninguna razón para no robar, y podría haber tenido otros deseos sin dejar de ser perfectamente racional.

La teoría humeana de las razones prácticas es mi principal oponente. Pero la teoría humeana no se apoya en buenos argumentos y es inverosímil en su tratamiento incluso de los casos que no implican moralidad. Argumentaré que los humeanos no pueden ofrecer una comprensión plausible del fenómeno de la debilidad de la voluntad, ni pueden explicar por qué es racional renunciar a las satisfacciones inmediatas en aras de mayores beneficios a largo plazo, ni pueden explicar de forma plausible por qué la moralidad tiene la importancia que tiene para nosotros y por qué es racional actuar moralmente.

Explicaré y defenderé todos estos puntos en los capítulos siguientes. Pero muchos filósofos consideran que el intuicionismo es increíble a primera vista. Como dice David Lewis, no puedo refutar una mirada incrédula;[11] sin embargo, me enfrentaré a todas las principales objeciones planteadas contra el intuicionismo. Al examinarlas, estas objeciones resultan mucho más débiles de lo que generalmente se supone. Por ejemplo, la "refutación" más popular del intuicionismo cita la prevalencia de las disputas morales como prueba en contra de la objetividad o la posibilidad de conocer las afirmaciones morales. La

suposición parece ser que si hubiera verdades morales objetivas y tuviéramos una manera de conocerlas, entonces habría poco desacuerdo moral. Sin embargo, nuestra experiencia con otros asuntos cuya objetividad es difícilmente discutible, no proporciona ningún apoyo para esa predicción. Lamentablemente, hay bastante controversia sobre la eficacia de la astrología, sobre si las leyes de control de armas reducen la delincuencia y sobre quién es probable que gane la próxima Super Bowl; sin embargo, casi nadie niega que estas cuestiones sean "objetivas" o que sea posible tener creencias justificadas sobre ellas. Los defensores del argumento del desacuerdo, como yo lo llamo, tampoco tienen una explicación preparada de por qué los mecanismos que conducen a los desacuerdos sobre estas cuestiones no deberían funcionar en los asuntos morales. De hecho, estos defensores parecen operar con una imagen simplista de la psicología humana en la que la irrealidad de un tema es la única explicación posible para el fracaso del consenso sobre el mismo.

1.6 Supuestos previos

Queda un asunto preliminar que discutir antes de proceder al caso del intuicionismo ético. Se trata de los supuestos filosóficos generales que doy por sentados en este libro.

Imaginemos a un par de científicos debatiendo los méritos de la Teoría General de la Relatividad. El científico *A* cita la observación de Eddington en 1919 de la curvatura gravitatoria de la luz alrededor del sol como prueba a favor de la teoría. El científico *B* pregunta entonces cómo sabía Eddington que no estaba soñando o, de hecho, cómo sabemos que los sentidos son una fuente fiable de información sobre la realidad externa. ¿Tiene *A* que responder a esto?

No; las preguntas de *B* no son un movimiento justo en un debate

sobre física. Una de las razones es pragmática: si aceptamos las exigencias dialécticas de *B*, entonces casi todas las discusiones pueden descarrilar en un debate sobre el escepticismo filosófico. No es que considere que esa cuestión no merezca ser discutida; de hecho, en otro momento, estaría encantado de discutirla.[12] Pero una discusión sobre física no es el lugar para plantearla. Si a uno le preocupan las consideraciones escépticas generales, debería hablar de ello con un epistemólogo, e incluso sólo en determinados contextos de conversación. Por supuesto, si el escepticismo es realmente cierto, entonces la relatividad general no está justificada, ni por las observaciones de Eddington ni por ninguna otra cosa. Pero cuando hablamos de física, asumimos tácitamente la falsedad del escepticismo. Sin eso, nunca avanzaríamos.

Más allá de las consideraciones pragmáticas, se han articulado y defendido muchas respuestas al escepticismo filosófico. Es razonable, incluso sin haber examinado la cuestión en detalle, pensar que alguna respuesta al escepticismo en epistemología es correcta, una que nos permita tener al menos creencias razonables sobre el mundo físico, y probablemente no importa a efectos de hacer física cuál es exactamente esa respuesta.

Espero que el lector esté de acuerdo hasta aquí. Después de todo, a casi nadie en la vida real se le ocurriría argumentar de la forma en que me imaginé al científico *B*. Una regla similar se aplica en las discusiones metaéticas. Por ejemplo, no deberíamos considerar justo que alguien que argumente contra el intuicionismo ético despliegue argumentos escépticos generales. En otras palabras, en esta discusión, puedo asumir que estamos justificados en creer la mayoría de las cosas que normalmente consideramos justificadas en creer, fuera del área de la ética. Puedo suponer que está justificado creer que las rocas existen, que

somos conscientes, que la Tierra seguirá existiendo la próxima semana, que 4+1=5 y que ningún objeto puede ser completamente rojo y completamente azul. Por lo tanto, si se puede demostrar que algún argumento particular contra el intuicionismo es simplemente un caso especial de un argumento más general que impugna nuestro conocimiento de ese tipo de cosas, entonces puedo dejar de lado ese argumento por no ser relevante para la discusión actual.

Lo que vale para el escepticismo filosófico general vale también para el antirrealismo ("metafísico") general. Es decir, al igual que no necesito, al defender la posibilidad del conocimiento moral, responder a los argumentos contra la posibilidad del conocimiento en general, no necesito, al defender la objetividad del valor, responder a los argumentos contra la posibilidad de la realidad objetiva en general.

Los lectores aquejados de sentimientos escépticos o antirrealistas generales pueden, pues, considerar las tesis principales de este libro como condicionales: si hay, en general, una realidad objetiva, y podemos saber cosas sobre ella, entonces hay también verdades evaluativas objetivas, de las que podemos saber algo. Es decir, no hay ningún problema especial para la objetividad y el conocimiento moral.

Notas

1. Algunas precisiones sobre el uso que hago de "enunciado evaluativo": En primer lugar, utilizo el término de forma amplia, de modo que incluye tanto los enunciados deónticos como los axiográficos. En segundo lugar, "enunciado evaluativo" se refiere a los tipos de frases que la mayoría de la gente consideraría como enunciados evaluativos, como "El placer es bueno" y "Barney es malo", independientemente de que (como discuten los no cognitivistas) esas cosas expresen proposiciones genuinas. En tercer lugar, también contaré esas frases como "enunciados evaluativos" cuando aparezcan incrustadas en oraciones más amplias; así, en "Si el placer es bueno, entonces la heroína es buena", llamo a la frase "el placer es bueno" un enunciado evaluativo, aunque no esté,

estrictamente hablando, funcionando como un enunciado allí. En cuarto lugar, incluyo como "evaluativos" los enunciados que aplican predicados evaluativos gruesos, como "justo" o "vicioso", cuyo significado incluye una evaluación además de una afirmación descriptiva.

2. Haidt, Koller y Díaz (1993) han discutido la putativa maldad de comerse al perro de la familia después de que haya sido atropellado por un coche.

3. Esto es utilizar "conciencia" en su sentido relacional, aquel en el que se habla de conciencia de algo. Véase mi (2001, pp. S 1-7) para más información sobre el concepto de conciencia.

4. Compárese la observación de Hume ([1739]1992, p. 469) de que la viciosidad de una acción "reside en uno mismo, no en el objeto". Véase Fumerton (1990, pp. 35-8) para una definición similar.

5. Donde "F" es cualquier predicado. Esta definición pretende captar la forma en que se suele utilizar la palabra en los debates sobre la "objetividad" del valor, aunque no todo el mundo la utiliza así. Hare (1999, p. 11) dice que un principio moral objetivo sería aquel "con el que ninguna persona racional que conozca los hechos podría estar en desacuerdo", y Adams (1981, p. 91) dice que un hecho objetivo es aquel que no depende de que un ser humano piense que lo tiene. La definición de Hare me parece demasiado estrecha (algunos hechos físicos podrían resultar "subjetivos" según su punto de vista), mientras que la de Adams parece demasiado amplia (parece hacer que todo sea "objetivo").

Hay que reconocer que mi definición no lo dice todo. Es más probable que llamemos subjetivo a "Jon Stewart es divertido" que a "Jon Stewart está bien", aunque ambas frases describen las actitudes de la gente hacia Jon Stewart. Tal vez la diferencia sea que sólo la primera frase parece atribuir una propiedad no relacional a Stewart; como "le gusta" es el participio pasado de "gustar", la segunda frase se refiere explícitamente a una actitud hacia Stewart.

6. Aquí y en todo el texto, distingo entre "supuesta referencia" y "referencia exitosa" (o simplemente "referencia"). Un predicado "F" se refiere supuestamente a un tipo de propiedad si los enunciados de la forma "a es F" significan que, o se usarían normalmente para afirmar que, a tiene una propiedad de ese tipo. "F" (con éxito) se refiere a un tipo de propiedad si supuestamente se refiere a una propiedad de ese tipo y de hecho hay una única propiedad de ese tipo que satisface el significado del predicado.

7. Veremos excepciones a esto en el capítulo 2, cuando lleguemos a discutir los puntos

de vista de Hare y Blackburn. Pero la presente caracterización del no-cognitivismo será suficiente por ahora.

8. Es decir, afirmaciones evaluativas que implican que algo tiene una propiedad evaluativa. Un nihilista puede aceptar la verdad de enunciados aparentemente evaluativos como "No es cierto que quemar gatos esté mal" y "Quemar gatos está mal, o no lo está".

9. También existe la posibilidad de que uno de estos puntos de vista sea válido para algunos enunciados evaluativos, mientras que otro lo sea para otros. Y en un punto de vista subjetivista, sin duda habrá algunos enunciados morales falsos (como "Comer bebés es obligatorio")", de modo que se podría decir que el nihilismo es cierto para estos.

10. En esta discusión, "el mundo" significa el mundo aparte de nuestro lenguaje y nuestros conceptos; del mismo modo, "hechos no valorativos" significa hechos no valorativos distintos de los del lenguaje y los conceptos. Por lo tanto, las disputas sobre "el mundo" y "los hechos no valorativos" se contradicen con las disputas semánticas.

11. Lewis (1986, p. 133).

12. Por "escepticismo", me refiero a la opinión de que no tenemos creencias justificadas sobre el mundo externo, o quizás sobre nada en absoluto. Véase mi (2001) para un tratamiento extenso de este punto de vista.

Parte I:
Teorías Metaéticas Alternativas

2

El no cognitivismo

2.1 No cognitivismo clásico

El placer es bueno.

¿Cómo debemos entender esta afirmación? La respuesta más directa es la cognitivista. El cognitivismo ético es la opinión de que los enunciados evaluativos como "El placer es bueno" afirman proposiciones que pueden ser verdaderas o falsas, igual que los enunciados "El cielo es rojo" y "Las comadrejas son mamíferos". Teniendo esto en cuenta, la explicación más directa de lo que hace la palabra "bueno" en la frase es la siguiente: hay una propiedad, la bondad, a la que se refiere la palabra, y la frase atribuye esa propiedad al placer.[1]

Los no cognitivistas niegan que "bueno" denote una propiedad, y niegan que "El placer es bueno" afirme nada del modo en que lo hace "Las comadrejas son mamíferos". Por tanto, les corresponde a ellos darnos una explicación alternativa del significado de "El placer es bueno". Han surgido dos enfoques no cognitivistas principales: el emotivismo y el prescriptivismo.

El emotivismo sostiene que la función principal de "*x* es bueno" es

expresar una emoción o actitud positiva (no una creencia) que el hablante tiene sobre *x*. Así, "*x* es bueno" es como "¡Hurra por *x*!". Cuando se dice: "El placer es bueno", no se está afirmando que se tenga ese sentimiento; sólo se expresa el sentimiento. Compare: cuando uno dice "¡Ay!", no está afirmando que siente dolor; sólo está expresando su dolor. Una diferencia entre afirmación y expresión es que una afirmación, como "Me duele algo", es verdadera o falsa, mientras que una mera expresión, como "¡Ay!", no es ni verdadera ni falsa. Así, los emotivistas han negado tradicionalmente que "*x* es bueno" sea verdadero o falso.[2]

El prescriptivismo sostiene que la función de "*x* es bueno" es, a grandes rasgos, decirle a la gente lo que tiene que hacer. Obsérvese que "Perseguir *x*!" no es una afirmación y no es ni verdadero ni falso. Del mismo modo, "*x* es bueno" no es ni verdadero ni falso desde un punto de vista prescriptivista simple.[3] Es cierto que los enunciados morales difieren de otros imperativos en algunos aspectos importantes, por lo que "*x* es bueno" no es exactamente igual que "Persigue *x*". Por ejemplo, parece que los imperativos morales tienen una especie de universalidad: al plantear un imperativo moral, uno respalda una especie de regla general aplicable a todas las personas, incluido uno mismo, lo que no ocurre con los imperativos ordinarios no morales".[4]

Existen otras formas más sutiles de no cognitivismo. Quizá las afirmaciones evaluativas tengan funciones tanto emotivas como prescriptivas. Quizá sirvan tanto para influir en los sentimientos del público como para expresar los del orador.[5] Quizá expresen algún otro estado mental no cognitivo, además de la emoción y el deseo.[6] Lo que es común a todos los puntos de vista no cognitivistas es, de nuevo, que sostienen que la función principal de los enunciados evaluativos es algo distinto de la afirmación de proposiciones.[7] Los argumentos de este

capítulo apoyarán el cognitivismo contra cualquier forma de no cognitivismo.

2.2 ¿Cómo podemos saber si el cognitivismo es cierto?

Comienzo con una observación sencilla pero crucial para el debate sobre el cognitivismo ético. La observación es que la cuestión que nos ocupa es empírica. El cognitivismo y el no cognitivismo son afirmaciones contrapuestas sobre el significado y la función actual del lenguaje evaluativo. Los no cognitivistas no abogan por un cambio en la forma en que utilizamos el lenguaje, lo que equivaldría a un esfuerzo por cambiar de tema. Intentan describir cómo funciona realmente el lenguaje evaluativo. Y este tipo de afirmaciones son empíricas. Imaginemos que alguien sostiene que la palabra "pato" significa "animal de cuatro patas". Esta hipótesis tiene predicciones comprobables: por ejemplo, que los angloparlantes típicos (competentes en el uso de estas palabras) no llamarían a nada "animal de cuatro patas" a menos que también lo llamaran "pato". Podemos comprobar esta predicción, por ejemplo, preguntando a algunos hablantes competentes si llamarían "animal de cuatro patas" a un gato y también si lo llamarían "pato".

Es un ejemplo sencillo. Refutar el no cognitivismo no será tan sencillo, porque el no cognitivismo es una tesis más abstracta que no tiene por qué implicar la afirmación de que "bueno" es sinónimo de cualquier otra expresión. Sin embargo, la cuestión sigue siendo que el cognitivismo y el no cognitivismo, como afirmaciones sobre significados, deberían tener predicciones comprobables sobre la forma en que utilizamos el lenguaje evaluativo. Podemos ver si la gente usa los enunciados evaluativos más como oraciones típicas que expresan proposiciones, o más como oraciones típicas que no expresan

proposiciones. Lo veremos en la siguiente sección.

No todo el mundo plantea la cuestión de esta manera. Y hay que admitir que existen otras consideraciones relevantes para evaluar el no cognitivismo ético además de si tratamos las afirmaciones éticas como aserciones. Normalmente, los filósofos se basan en consideraciones teóricas más amplias, como la supuesta conveniencia, en general, de mantener una "visión naturalista del mundo" (evitando las cosas "raras" de las que no hablan los científicos), la supuesta insostenibilidad de las teorías metaéticas alternativas y la capacidad del no cognitivismo para explicar por qué estamos motivados para actuar de la forma que consideramos moral.

La apelación a estas amplias consideraciones teóricas es metodológicamente inferior al tipo de apelación que haré a cómo utilizamos el lenguaje ético. ¿Por qué? Porque las pruebas lingüísticas son, con diferencia, las más objetivas. Aquí utilizo "objetiva" en un sentido epistemológico (no en el sentido tratado en la sección 1.1): una prueba objetiva es, a grandes rasgos, una prueba cuya existencia y relevancia pueden determinarse de un modo relativamente sencillo; no requiere difíciles ejercicios de juicio, sopesar consideraciones contrapuestas, resolver complejas cuestiones filosóficas o establecer una visión global del mundo para determinar si la prueba existe o qué conclusión apoya.

Por ejemplo, A.J. Ayer defendió el no cognitivismo apelando a las premisas (1) de que los enunciados éticos son inverificables, y (2) de que un enunciado sólo tiene significado cognitivo si puede verificarse. En resumen, Ayer citó la inverificabilidad de los enunciados éticos como prueba contra el cognitivismo.[8] Esta prueba es muy subjetiva, ya que para determinar si los enunciados éticos son realmente inverificables, primero debemos resolver cuestiones filosóficas como si la intuición

cuenta como medio de conocimiento, si los hechos sobre nuestras actitudes y convenciones cuentan como verificación para los enunciados evaluativos, y si los enunciados evaluativos ayudan alguna vez a explicar hechos no evaluativos. Además, para aceptar la inverificabilidad de los enunciados evaluativos como prueba del no cognitivismo, primero hay que aceptar la filosofía positivista lógica de Ayer sobre lo que es el "significado". Todas estas son cuestiones filosóficas muy polémicas y complejas. Lo que quiero decir no es que estas cuestiones no puedan resolverse, ni que los argumentos que implican afirmaciones filosóficas controvertidas nunca sean apropiados. Lo que quiero decir es que sería una tontería confiar en pruebas de ese tipo cuando se dispone de pruebas directas y objetivas de una tesis.

Mi metodología es más o menos la siguiente. Hay algunas frases que son incontrovertiblemente asertivas, frases que tanto los cognitivistas éticos como los no cognitivistas considerarían que afirman proposiciones. Hay otras oraciones que son incontrovertiblemente no asertivas. Podemos identificar características de las primeras que no pertenecen a las segundas, y ver si los enunciados evaluativos tienen esas características. Las pruebas que voy a citar están al alcance de cualquier angloparlante competente. Tampoco es fácil discutir su pertinencia. Sin duda, algunos discutirán si las pruebas son suficientes para rechazar el no cognitivismo, pero incluso los no cognitivistas admiten que van en contra de su teoría.[9]

2.3 Las pruebas lingüísticas del cognitivismo

Hay varios rasgos característicos de las oraciones que expresan proposiciones, todos los cuales tienen los enunciados evaluativos:[10]

(a) Los enunciados evaluativos adoptan la forma de oraciones

declarativas, en lugar de, por ejemplo, imperativos, preguntas o interjecciones. "El placer es bueno" tiene la misma forma gramatical que "Las comadrejas son mamíferos". Las oraciones de esta forma se utilizan normalmente para hacer afirmaciones fácticas.[11] En cambio, los paradigmas de las expresiones no cognitivas, como "Hurra por *x*" y "Persigue a *x*", no son oraciones declarativas.

(b) Los predicados morales pueden transformarse en sustantivos abstractos, lo que sugiere que pretenden referirse a propiedades; hablamos de "bondad", "rectitud", etc., como en "No cuestiono la prudencia del acto, sino su rectitud".

(c) Atribuimos a las evaluaciones el mismo tipo de propiedades que a otras proposiciones. Se puede decir: "Es cierto que he hecho algunas cosas mal en el pasado", "Es falso que la anticoncepción sea un asesinato" y "Es posible que el aborto esté mal". "Verdadero", "falso" y "posible" son predicados que sólo se aplican a las proposiciones. Nadie diría: "Es verdad que ouch", "Es falso que cierra la puerta" o "Es posible que hurra".

(d) Todos los verbos de actitud proposicionales pueden anteponerse a enunciados evaluativos. Podemos decir: "Jon cree que la guerra fue justa", "Espero haber hecho lo correcto", "Ojalá tuviéramos un presidente mejor" y "Me pregunto si hice lo correcto". En cambio, nadie diría: "Jon cree que ouch", "espero que hurra por los Broncos", "ojalá cierra la puerta", o "me pregunto si por favor pasa la sal". La explicación obvia es que estados mentales como creer, esperar, desear y preguntarse son por naturaleza proposicionales: Esperar es tener la esperanza de que algo ocurra, preguntarse es preguntarse si algo ocurre, etcétera. Por eso no se puede esperar que uno haya hecho lo correcto a menos que exista una proposición -algo que podría ser el caso- que corresponda a la expresión "uno hizo lo correcto".

(e) Los enunciados evaluativos pueden transformarse en preguntas sí/no: Se puede afirmar "El helado de canela es bueno", pero también se puede preguntar "¿Es bueno el helado de canela?". No se pueden formar preguntas análogas a partir de imperativos o expresiones emocionales: "Cierra la puerta" y "Hurra por los Broncos" carecen de significado claro. La explicación obvia es que una pregunta sí/no requiere una proposición; pregunta si algo es el caso.

Un prescriptivista no cognitivista podría interpretar algunas preguntas evaluativas sí/no como peticiones de instrucciones, como en "¿Debo apagar el horno ahora?". Pero otras preguntas desafiarían esa interpretación, incluidas las preguntas evaluativas sobre el comportamiento de otras personas o sobre el pasado: "¿Estuvo mal que el emperador Nerón matara a Agripina?" no es una petición de instrucción.

(f) Uno puede emitir imperativos y expresiones emocionales dirigidas a cosas que se caracterizan moralmente. Si el no cognitivismo es cierto, ¿qué significan éstos? "Haz lo correcto". "Viva la virtud". Aún más desconcertante para el no cognitivista es imaginar contextos apropiados para comentarios como: "No deberíamos hacer esto, pero no me importa; hagámoslo de todos modos". Esto es perfectamente inteligible, pero sería ininteligible si "No deberíamos hacer esto" expresara una emoción aversiva hacia la acción propuesta o emitiera un imperativo de no hacerla.

(g) En algunas oraciones, los términos evaluativos aparecen sin que el hablante apoye o impugne nada, pero los términos se utilizan en su sentido normal. Esto se conoce como el problema Frege-Geach y constituye la base de la objeción más conocida al no cognitivismo.[12] Consideremos dos ejemplos.

Primer ejemplo:

> "Si el adulterio está mal, entonces Dios castigará a Clinton". Esta frase no emite ningún veredicto sobre el adulterio; no dice ni que el adulterio sea malo ni que no lo sea. Por lo tanto, "incorrecto" no se utiliza aquí para expresar una reacción psicológica negativa ante el adulterio, ni para emitir un imperativo contra el adulterio, ni para influir en los sentimientos de la audiencia contra el adulterio. Es difícil ver cómo el no cognitivista podría explicar el significado de "el adulterio está mal" en este contexto. De hecho, dado que un enunciado de la forma "Si *p* entonces *q*" generalmente requiere que se introduzca una proposición para *p*, parece que si "el adulterio es malo" no expresa ninguna proposición, entonces toda la frase no tiene sentido, como "Si por favor pasa la sal, entonces las patatas necesitan más sal" o "Si hurra por los Broncos, entonces los Broncos van a ganar".[13]

Segundo ejemplo:

> A bordo del Titanic que se hunde, el capitán dice: "Hay una forma correcta de manejar esta situación". Como no dice cuál es la forma correcta de manejar la situación, no está expresando una emoción positiva hacia nada, ni intentando influir en las actitudes de nadie, ni diciendo a nadie que haga nada. Sin embargo, es evidente que "correcto" se utiliza en su sentido normal. Esto se demuestra por el hecho de que el capitán podría continuar sensatamente de la siguiente manera: "Y esa manera es echar a suertes quién puede subir a los botes salvavidas". Aunque ahora ha apoyado la idea de echar a suertes, obviamente no ha cambiado el significado de su uso inicial de "correcto".

Una opción tentadora es dejar que los términos evaluativos se refieran a emociones o actos de dar órdenes cuando aparecen en estos contextos. Por ejemplo, podríamos leer "Si el adulterio está mal, entonces Dios

castigará a Clinton" como algo que significa "Si el adulterio entra en conflicto con los mandamientos de Dios, entonces Dios castigará a Clinton". He aquí por qué eso no ayudará al no cognitivista: ¿Significa la frase "El adulterio es malo", en sí misma, que el adulterio es contrario a la voluntad de Dios?

Supongamos que el no cognitivista dice "sí". Entonces "El adulterio está mal" expresa una proposición, que es verdadera o falsa: es verdadero o falso que el adulterio entra en conflicto con los mandamientos de Dios. Nuestro antiguo no-cognitivista ha abandonado así el no-cognitivismo en favor del subjetivismo: son los subjetivistas los que dicen que los enunciados evaluativos expresan proposiciones sobre las actitudes de alguien. Criticaremos el subjetivismo en el capítulo siguiente; por ahora, lo importante es que ésta no es una manera de preservar el no-cognitivismo.

Supongamos, entonces, que el no cognitivista dice "no". Consideremos ahora la siguiente inferencia deductiva que alguien podría hacer:

> Si el adulterio está mal, Dios castigará a Clinton.
>
> El adulterio está mal.
>
> Por lo tanto, Dios castigará a Clinton.

Nuestro no cognitivista se compromete ahora a afirmar que "el adulterio está mal" tiene un significado distinto en la primera premisa que en la segunda. Pues ha dicho que en la primera premisa, "el adulterio está mal" significa que el adulterio entra en conflicto con los mandamientos de Dios. Pero en la segunda premisa no significa esto; en su lugar, sólo expresa alguna actitud no cognitiva. De ser así, el argumento anterior debería parecernos ilógico, como si cometiera un equívoco. Pero no es así: podemos ver que las dos primeras afirmaciones implican la tercera.

Esto sólo puede ser así si "erróneo" tiene el mismo significado en ambas premisas.[14]

Lo mismo se aplica a cualquier otro intento de salvar el no cognitivismo dejando que los términos evaluativos se refieran a las emociones u órdenes de alguien. Así, supongamos que el no cognitivista dice que "Hay una manera correcta de manejar esta situación" significa "Hay una manera de manejar esta situación que yo aprobaría". Ahora pregúntese: ¿"*x* es correcto", por sí mismo, significa "yo aprobaría *x*"? Si el no cognitivista dice "sí", entonces ha abandonado el no cognitivismo en favor del subjetivismo. "Yo aprobaría *x*" es una afirmación de hecho, que puede ser verdadera o falsa, no un enunciado no cognitivo.

Si el no cognitivista dice "no", entonces debe decir que "correcto" cambia de significado entre las dos frases siguientes:

> Hay una manera correcta de manejar esta situación.
>
> La manera correcta de manejar esta situación es echar a suertes quién sube a los botes salvavidas.

-ya que "manera correcta" en la primera frase significa "manera que yo aprobaría", pero "manera correcta" en la segunda frase no significa esto, sino que funciona para expresar una actitud emocional no cognitiva hacia el sorteo de pajitas. Si es así, entonces la segunda afirmación no implica la primera. Pero eso es falso: "correcto" tiene obviamente el mismo significado en ambas frases, y la segunda frase obviamente implica a la primera.

Resumamos los argumentos expuestos en los puntos (b)-(g) anteriores. He aquí una lectura fuerte del argumento contra el no cognitivismo:

1. Cada una de las siguientes frases tiene sentido:

Cuestiono la corrección del acto.

Es cierto que el placer es bueno.

Espero haber hecho lo correcto.

¿Está mal el aborto?

Haz lo correcto.

Si el placer es bueno, entonces el chocolate es bueno.

Algo es bueno.

2. Ninguna de esas frases tendría sentido si el no cognitivismo fuera cierto.

3. Por lo tanto, el no cognitivismo es falso.

Cada una de las frases enumeradas en la premisa (1) crea una objeción distinta al no cognitivismo; el no cognitivista debe tratar todos los ejemplos para defender su posición.

La justificación de la premisa (1) es que las oraciones enumeradas, y otras similares, tendrían sentido para cualquier angloparlante normal; de hecho, no son ni siquiera ligeramente extrañas. Casi todo el mundo las consideraría indiscutiblemente ciertas. Prácticamente nadie discutiría la veracidad de mi afirmación de que creo que el placer es bueno, o que Si el aborto está mal, entonces los fetos tienen derechos. El argumento *prima facie* para considerar inteligible cada una de esas frases es tan sólido como para considerar inteligible cualquier cosa.

La premisa (2) parece plausible si se tiene en cuenta la tesis no cognitivista de que "El placer es bueno" significa algo así como "Viva el placer" o "Persigue el placer".

(1). En algunos casos (como en "Espero haber hecho lo correcto"), es

difícil siquiera pensar en una paráfrasis de las frases basada en la teoría no cognitivista. En otros casos, se puede pensar en una paráfrasis de primer paso, pero evidentemente es una frase sin sentido (como "Si hurra por el placer, entonces hurra por el chocolate").

(3) se deduce de (1) y (2). Supongo que el no cognitivista querría atacar la premisa (2), ya que atacar (1) parece inútil. Consideraremos los ataques a la premisa (2), al menos para algunos de los tipos de oraciones enumerados, en las siguientes secciones de este capítulo.

He aquí una interpretación más débil del problema para el no cognitivismo. En esta interpretación, el argumento no refuta de manera concluyente el no cognitivismo; sin embargo, tiene fuerza contra el no cognitivismo incluso si el no cognitivista puede encontrar una manera de dar sentido a las oraciones enumeradas en la premisa (1) anterior.

4. Si el no cognitivismo fuera cierto, entonces sería más probable que las siguientes oraciones nos parecieran extrañas, malformadas o confusas, que si el cognitivismo fuera cierto:

    Cuestiono la corrección del acto.

    Es cierto que el placer es bueno.

    Espero haber hecho lo correcto.

    ¿Está mal el aborto?

    Haz lo correcto.

    Si el placer es bueno, entonces el chocolate es bueno.

    Algo es bueno.

5. Ninguna de esas frases nos parece extraña, malformada o confusa.

6. Por lo tanto, nuestra reacción a esas frases es una prueba a favor del

cognitivismo sobre el no cognitivismo.

Podemos ver que (4) es cierto contrastando enunciados paradigmáticamente no cognitivos. No hay, que yo sepa, ninguna expresión u oración no cognitiva no controvertida a partir de la cual se puedan formar oraciones análogas a cualquiera de las enumeradas en (4). "Hurra por los Broncos" y "Pásame la sal" no dan lugar a oraciones como ésas. Así que al menos debemos decir que, si el no cognitivismo no implica que las oraciones enumeradas nos parezcan ininteligibles, al menos nos lleva a esperar que así sea. Imaginemos cómo serían las cosas si las frases enumeradas nos parecieran extrañas, malformadas o confusas. Si a la mayoría de los angloparlantes competentes les pareciera ininteligible "Creo que el placer es bueno"/"I believe pleasure is good", igual que a nosotros nos parece ininteligible "Creo que por favor pasa la sal"/"I believe please pass the salt", es evidente que los no cognitivistas podrían citarlo como una prueba contundente de su tesis. Por el contrario, si no nos parece ininteligible "creo que el placer es bueno", entonces este hecho es una prueba en contra del no cognitivismo.[15]

Tampoco hay que subestimar la fuerza de estas pruebas. El no cognitivista hace una afirmación sobre significados reales, pero todas las pruebas razonablemente directas y objetivas, extraídas de la forma en que utilizamos el lenguaje ético, apuntan hacia el cognitivismo. Cada prueba lingüística que diferenciaría ejemplos incontrovertibles de enunciados cognitivos de ejemplos incontrovertibles de enunciados no cognitivos arroja el resultado de que los enunciados evaluativos encajan con los cognitivos. Los enunciados evaluativos actúan en todos los sentidos como afirmaciones fácticas.

Hasta ahora, los no cognitivistas han apoyado sus argumentos en el tipo de argumentos filosóficos generales que, según he argumentado anteriormente, son metodológicamente inferiores a la confianza en las

intuiciones lingüísticas y los hechos sobre el uso del lenguaje. A pesar de mi objeción metodológica, no tengo intención de ignorar sus argumentos; los retomaré en capítulos posteriores. En general, pueden considerarse argumentos contra la viabilidad de otras posturas metaéticas, en particular el realismo moral. Si las demás posturas metaéticas resultaran realmente inviables, entonces tendríamos motivos para volver a considerar el no cognitivismo. Pero por ahora deberíamos rechazarlo.

*2.4 Hare sobre la verdad moral

La mayoría de los argumentos expuestos en la sección anterior son conocidos de alguna forma por los filósofos desde hace más de 40 años, y sin embargo sigue habiendo filósofos no cognitivistas. En esta sección y en las dos siguientes, analizaré cómo responden algunos destacados no cognitivistas a algunos de estos problemas.

Para empezar, R. M. Hare ha abordado el problema de la verdad moral, esforzándose por demostrar que su versión del prescriptivismo puede preservar una noción de verdad y falsedad para los enunciados morales. ¿Cómo? En su trabajo se pueden encontrar dos propuestas distintas. En primer lugar, Hare propone que los enunciados morales tienen un significado descriptivo, de expresión de proposiciones, además de su significado prescriptivo. Es este significado descriptivo el que les permite ser verdaderos o falsos. Hare parece pensar que el significado descriptivo de un enunciado moral es equivalente al de un enunciado de los hechos no morales que son las razones del hablante para emitir el juicio moral.[16] Por ejemplo, supongamos que Jon tiene una teoría moral exhaustiva, según la cual la acción moralmente correcta en cualquier situación es siempre la acción que produce la mayor cantidad de

felicidad en el mundo. Jon dice: "Comer carne está mal". Según Hare, la afirmación de Jon tiene un significado descriptivo y otro prescriptivo. El significado descriptivo es: "Comer carne disminuye la felicidad en el mundo". El significado prescriptivo es algo así como "No comas carne".[17] La afirmación original, "Comer carne está mal", puede considerarse verdadera o falsa en función de si comer carne disminuye de hecho la felicidad en el mundo.

Creo que se trata de una falsa concepción del significado. Cuando uno hace una afirmación moral, el significado de la afirmación es una cosa; la razón que uno tiene para hacerla es otra, y esta última generalmente no forma parte de la primera. Por ejemplo, supongamos que George está a favor de invadir Irak. Su razón es que cree que Irak tiene armas de destrucción masiva, y piensa que esto es una justificación adecuada para la invasión.[18] Un segundo individuo, Tony, también cree que deberíamos invadir Irak. Pero Tony no cree que Irak tenga armas de destrucción masiva. La razón de Tony para estar a favor de la invasión es que Saddam Hussein es un dictador cruel que ha causado mucho sufrimiento al pueblo de Irak. George y Tony tienen el siguiente intercambio:

George: Deberíamos invadir Irak.

Tony: Estoy de acuerdo. Deberíamos invadir Irak.

En opinión de Hare, Tony está confundido al pensar que está de acuerdo con lo que ha dicho George. Porque aunque se pueda decir que Tony está de acuerdo con el plan de acción de George, parte del significado de la afirmación de George era que Irak tiene armas de destrucción masiva, cosa que Tony no cree. Así que Tony no está de acuerdo con la afirmación en su conjunto. Pero esto es obviamente falso. Aunque George puede haber estado pensando que Irak tiene armas de destrucción masiva, y aunque esto puede ser parte de la razón por la que

dijo lo que dijo, George, en el intercambio anterior, no dijo realmente que Irak tiene armas de destrucción masiva, ni dijo nada que implique esto.

No sé si Hare consideraría realmente que "Irak tiene armas de destrucción masiva" forma parte del significado de la declaración de George. Tal vez Hare nos pediría, en cambio, que formuláramos la teoría ética general y exhaustiva de George. Por desgracia, la mayoría de la gente no tiene esa teoría. Pero supongamos que en este caso tenemos suerte y George tiene una teoría ética global. Digamos que su teoría es el utilitarismo. Entonces, Hare podría decir que el significado descriptivo de la afirmación de George es: "Invadir Irak maximizaría la felicidad total en el mundo". Aunque esto pueda parecer menos absurdo que la sugerencia anterior, sigue siendo falso: al decir simplemente que deberíamos invadir Irak, George no dice que hacerlo maximizaría la felicidad, ni dice nada que lo implique; simplemente no ha dicho nada, hasta ahora, sobre por qué cree que deberíamos invadir Irak. Y si Tony tuviera una teoría ética competidora, esto no le impediría estar de acuerdo con la conclusión de George diciendo: "Sí, deberíamos invadir Irak".

Consideremos un segundo ejemplo, aún más perjudicial para la concepción de la verdad moral de Hare. Como en el primer caso, George cree que deberíamos invadir Irak porque Irak tiene armas de destrucción masiva. Jean-Pierre también cree que Irak tiene armas de destrucción masiva y está de acuerdo con George en todas las demás cuestiones no morales. Pero Jean-Pierre y George tienen valores morales fundamentales diferentes, por lo que Jean-Pierre no considera que estos hechos justifiquen la invasión. George dice: "Deberíamos invadir Irak". Según Hare, Jean-Pierre debería considerar que esa afirmación es cierta.[19] Porque es el significado descriptivo de la afirmación lo que

determina la verdad o la falsedad, y Jean-Pierre está de acuerdo con George en cuestiones descriptivas.

Hasta ahora, he hablado como si Hare se basara en las normas morales del hablante para determinar el significado descriptivo de una afirmación moral. De hecho, sin embargo, Hare se remite a las normas "comúnmente aceptadas" "dentro de una determinada sociedad".[20] Así, si George y Tony pertenecen a la misma sociedad, acabarían (afortunadamente) queriendo decir lo mismo con "Deberíamos invadir Irak". Pero es bastante fácil revivir los problemas anteriores simplemente suponiendo que George pertenece a una sociedad diferente de la de Tony o Jean-Pierre. Y la dependencia de las normas socialmente aceptadas crea otro problema para el planteamiento de Hare: cómo tratar a las personas que no están de acuerdo con las normas morales de su sociedad. Desde el punto de vista de Hare, estas personas se verían obligadas a calificar de "verdaderos" los juicios morales realizados sobre la base de las normas morales imperantes. Y cuando esos disidentes morales llamaran "bueno" a algo, parte de lo que estarían diciendo sería que esa cosa satisfacía las normas morales de la sociedad. Parece claro que, si las normas morales de alguien intervienen en la determinación del significado descriptivo de una afirmación moral, deberían ser las normas del hablante.

Entiendo que el primer esfuerzo de Hare por asegurar la verdad y la falsedad de las afirmaciones morales fracasa. Es, de hecho, el producto de una falacia desgraciadamente común a la que los filósofos todavía no se han inmunizado del todo: la de confundir las creencias de fondo de un hablante con el significado de sus palabras. Un caso especial de esto es confundir las razones de una persona para decir algo con el significado de la declaración. Cuando expresamos un juicio moral, no estamos informando de las razones de hecho que tenemos para emitirlo.

Hare hace un segundo intento de asegurar la verdad moral, utilizando un análisis no cognitivista de "verdadero". Hare sugiere que, en algunos casos, podríamos llamar "verdadera" a una afirmación moral, lo que significaría simplemente respaldarla: "no será una afirmación de que la respaldamos, sino una expresión de respaldo".[21]

Uno de los problemas de esta sugerencia es que la palabra "true" no se utiliza así en español. En inglés, "true" sólo se aplica a (cosas que expresan) proposiciones. Por tanto, no puedo decir: "Es cierto que, por favor, pásame la sal", aunque apoye el imperativo de pasar la sal. No puedo decir: "Es cierto que hurra por los Broncos", aunque apoye el sentimiento que hay detrás de la frase "Hurra por los Broncos". Por tanto, Hare tendría que afirmar que los enunciados evaluativos funcionan de forma diferente a todos los enunciados no cognitivos paradigmáticos en el sentido de que, por alguna razón, se les puede aplicar "verdadero".

Además, la interpretación no cognitivista de "verdadero" propuesta por Hare se enfrenta a muchos de los mismos problemas que un análisis no cognitivista de "bueno". La frase "*x* es verdadero" lleva todas las marcas lingüísticas de una afirmación factual. Podemos hablar de "la proposición de que *x* es verdadera" y de "la verdad de mi afirmación"; podemos decir "creo que *x* es verdadera", "espero que *x* sea verdadera", "me pregunto si *x* es verdadera", "si *x* es verdadera, entonces también lo es *y*", "¿juras decir la verdad?" y "hay muchas palabras verdaderas dichas en broma".[22] Y las consideraciones que suelen aducirse en apoyo del no cognitivismo sobre el "bien" -como la supuesta inviabilidad de las teorías alternativas y el carácter orientador de la acción de los juicios morales- no ayudarán al no cognitivista sobre la "verdad".

Una propuesta mejor, quizá más cercana a la intención de Hare, es la teoría de la redundancia de la "verdad": "es cierto que" es simplemente redundante. Es cierto que "El placer es bueno" sólo significa "El placer

es bueno", igual que "Es cierto que las vacas tienen cuatro estómagos" sólo significa "Las vacas tienen cuatro estómagos". Esto nos da una razón de principio para sostener que "Es cierto que el placer es bueno" no es cognitivo y que "Es cierto que las vacas tienen cuatro estómagos" es cognitivo, sin tener que decir que "cierto" cambia de significado.

El problema con esta sugerencia es que no puede dar sentido a enunciados en los que se utiliza "verdadero" o "verdad" sin especificar lo que es o podría ser verdadero, como "¿Jura decir la verdad?" y "Muchas palabras verdaderas se dicen en broma". Ni la teoría de la redundancia ni la teoría de la aprobación de "verdadero" son válidas para estos ejemplos. Tampoco podemos interpretar estas afirmaciones en la línea de "¿Juras decir lo que apruebas?" y "Muchas palabras aprobadas por el hablante se dicen en broma", ya que cualquiera de estas interpretaciones abandonaría el tratamiento no cognitivista de la verdad por una teoría subjetivista, tratando la verdad meramente como la propiedad de ser aprobada por el hablante.[23] Obsérvese que las frases de este tipo pueden referirse tanto a verdades morales como no morales: "Jon dijo la verdad" puede decirse en el mismo sentido cuando la observación principal de Jon fue "Los gatos tienen dientes" que cuando su observación principal fue "Quemar gatos está mal".

Algunas de las propias afirmaciones de Hare constituyen contraejemplos del no cognitivismo tanto sobre la ética como sobre la verdad. Inmediatamente después de la propuesta de Hare de la teoría de la aprobación de la "verdad" moral, encontramos estas observaciones:

> Así que no podemos hacer que las normas sean aceptables simplemente aceptándolas; más bien, al decir que son aceptables estamos expresando nuestra aceptación de las mismas; y si debemos aceptarlas es una cuestión de pensamiento racional ....
>
> Tenemos que pensar, y hay maneras de hacerlo bien.

> Si lo hacemos bien, alcanzaremos el conocimiento de la verdad de las afirmaciones morales. . . .[24]

"Aceptable", "deberíamos", "tenemos que" y "bien" son expresiones evaluativas. No está claro qué se supone que significa cualquiera de las afirmaciones anteriores en una interpretación no cognitivista de las expresiones evaluativas. Tampoco puedo entender qué se supone que significa "Alcanzaremos el conocimiento de la verdad de los enunciados morales" si "verdad" es un término no cognitivo que sólo sirve para expresar la aprobación de un enunciado dado. Resulta tentador, sobre todo teniendo en cuenta el resto de las opiniones de Hare, interpretar que esto significa algo así como: "Llegaremos a saber qué enunciados morales aprobarían todos los seres racionales tras una reflexión suficiente". Pero aceptar esta interpretación supondría abandonar la teoría no cognitivista de la verdad propuesta en favor de una cognitivista en la que "verdadero", aplicado a enunciados evaluativos, significa "sería aprobado por todos los seres racionales tras una reflexión suficiente". Pero si esas son las condiciones de verdad de los enunciados evaluativos, entonces también podríamos analizar los términos morales como referidos a lo que los seres racionales aprobarían, y abandonar el no cognitivismo por el subjetivismo. Concluyo que Hare no ha logrado articular un sentido legítimo de la verdad moral coherente con su metaética.

*2.5 Los mundos factual-normativos de Gibbard

Lo último en tecnología antirrealista, los "mundos factual-normativos" de Allan Gibbard son objetos abstractos diseñados para ayudar a explicar los significados tanto de enunciados evaluativos simples como de oraciones complejas con componentes evaluativos. Un mundo

factual-normativo es una combinación de una especificación completa de cómo es el mundo, *w*, con un conjunto completo de reglas para actuar y para aprobar/desaprobar cosas, *n*. Podemos representar un mundo factual-normativo como un par ordenado, (*w*, *n*).[25]

El significado de un enunciado fáctico ordinario sin términos evaluativos puede representarse mediante un conjunto de mundos posibles, es decir, el conjunto de todas las formas en que podría ser el mundo para que el enunciado fuera verdadero.[26] Del mismo modo, Gibbard opina que podemos representar el significado de un enunciado que contenga términos evaluativos mediante un conjunto de mundos fáctico-normativos: el conjunto de todos los mundos fáctico-normativos en los que el enunciado es válido. Una afirmación es "válida" en un mundo factual-normativo (*w*, *n*) si, aproximadamente, alguien que pensara que el mundo es como w y que apoyara las normas n aceptaría la afirmación. Por ejemplo, supongamos que acabas de robar un caramelo a un niño y un amigo te dice: "Si robar ese caramelo estuvo mal, Dios te castigará". Ahora imagine dos mundos posibles, w_1 y *w2*, y dos posibles conjuntos completos de normas, n_1 y n_2, donde:

En w_1, Dios te castiga.

En *w2*, Dios no te castiga.

n_1 permite el robo de caramelos.

n_2 prohíbe el robo de caramelos.

A partir de estos mundos y estos conjuntos normativos, se pueden formar cuatro mundos factual-normativos: (w_1, n_1), (w_1, n_2), (*w2*, n_1) y (*w2*, n_2). La afirmación "Si robar el caramelo estuvo mal, entonces Dios te castigará" es válida en (w_1, n_1), (w_1, n_2) y (*w2*, n_1); no es válida en (*w2*, n_2).[27] Por tanto, (*w2*, n_2) es el único de los cuatro mundos que la afirmación del hablante excluye. El punto de vista de Gibbard puede

aplicarse igualmente bien a enunciados más complejos y a enunciados más generales sobre acciones ilícitas (por ejemplo, "Si está mal hacer algo, entonces está mal hacer que tu hermano pequeño lo haga").

¿Qué consigue el dispositivo de los mundos fáctico-normativos? Proporciona a Gibbard una forma de decir que el tipo de afirmaciones que nos preocupaban en la sección 2.3 -como "Si robar está mal, Dios te castigará"- tienen sentido. Es importante destacar que Gibbard no necesita plantear ningún cambio en el significado de "malo" de "Robar está mal" (afirmado por sí mismo) a "Si robar está mal, Dios te castigará". Tampoco tiene que dar teorías separadas de los significados de los distintos tipos de oraciones que contienen términos evaluativos; habrá reglas generales que relacionen los significados de las oraciones complejas con los de sus partes más simples, que son las mismas para los enunciados evaluativos y no evaluativos; por ejemplo, el no cognitivista no tendrá que decir que "si... entonces" tiene significados distintos dependiendo de si el antecedente es un enunciado evaluativo o descriptivo.

Para obtener estas atractivas ventajas, sólo tenemos que establecer el vínculo entre los mundos fáctico-normativos y los significados. ¿Qué tiene que ver exactamente un conjunto de mundos fáctico-normativos con lo que quiero decir cuando hago una afirmación que contiene "incorrecto"? Gibbard cree que explicar el significado de una afirmación es explicar el estado de ánimo que expresa. Su opinión no es, por supuesto, que yo tenga realmente en mente el conjunto relevante de mundos factual-normativos -los mundos posibles son típicamente demasiado complejos para que cualquier mente humana los conciba plenamente, por no hablar de los infinitos conjuntos de mundos factual-normativos a los que se refiere la teoría de Gibbard-. Más bien, la opinión de Gibbard es que tenemos dos tipos de estados mentales,

creencias y actitudes de aceptación de normas, y se puede decir que los estados mentales de cada uno de estos tipos descartan ciertos mundos factual-normativos. Esta "exclusión" es una cuestión de relaciones lógicas objetivas, de modo que mis creencias y normas pueden excluir cosas que no me doy cuenta de que excluyen (porque no he seguido todas las consecuencias lógicas de mis creencias y normas). He aquí un par de ejemplos: Si creo que mañana va a nevar, mi creencia excluye cualquier mundo factual-normativo en el que no nieve mañana. Si acepto la norma de no matar a la gente, entonces descarto cualquier mundo factual-normativo en el que se permita matar a personas. Obsérvese que esta norma no sólo descarta algunas normas competidoras, sino que también descarta algunas combinaciones de hecho-norma. Por ejemplo, mi norma de no matar personas excluye la siguiente combinación: la afirmación fáctica de que vender cigarrillos mata personas + la norma de aprobar la venta de cigarrillos. Por tanto, cualquier mundo fáctico-normativo que incluya ambas cosas quedaría descartado; sin embargo, un mundo fáctico-normativo que apruebe la venta de cigarrillos pero que contenga la afirmación fáctica de que vender cigarrillos no mata a la gente no quedaría descartado.

Cuando haces una afirmación, expresas un estado mental -un conjunto de creencias y normas aceptadas- que descarta determinados mundos fáctico-normativos. El conjunto de mundos fáctico-normativos en los que tu afirmación es válida es el conjunto de mundos que tu estado mental no descarta. Eso es lo que el conjunto de mundos que hemos estado discutiendo tiene que ver con el significado.

Puede parecer que todo esto resuelve el problema Frege-Geach. Gibbard parece creerlo así. Pero en realidad no toca el problema de fondo.

Empecemos con un experimento de filosofía del lenguaje. Está

comúnmente aceptado (aunque no universalmente) que hay cosas llamadas "proposiciones" sobre las que versan nuestras creencias. Si crees que lloverá mañana, tu creencia se dirige a la proposición de que lloverá mañana. Si crees que Snoopy se mojará, entonces tu creencia se dirige a la proposición de que Snoopy se mojará. Ahora imaginemos que nos encontramos con un filósofo que sostiene la interesante tesis de que no hay proposiciones compuestas. Por ejemplo, existe la proposición de que lloverá mañana y la proposición de que Snoopy se mojará, pero no existe la proposición de que si llueve mañana, Snoopy se mojará. Este filósofo tendría problemas, porque ciertamente existe la creencia de que si llueve mañana, Snoopy se mojará. El filósofo tendría dificultades para dar cuenta de esto, ya que no dispondría de ninguna proposición a la que pudiera dirigirse esta creencia. Admite el estado de creer que mañana lloverá, así como el estado de creer que Snoopy se mojará. Así que podría intentar construir de alguna manera la creencia compuesta a partir de esas dos creencias. Pero la persona que cree que si llueve mañana Snoopy se mojará, no necesita creer ni que lloverá mañana ni que Snoopy se mojará. Así que parece que el filósofo carecería de los recursos para hacer lo que necesita.

Consideremos ahora una creencia diferente (supuesta): la creencia de que si estuvo mal robar el caramelo, entonces Dios le castigará. Con respecto a esta creencia, el no cognitivista se encuentra en la misma posición que el filósofo hipotético del último párrafo. El no cognitivista reconoce un estado de aceptación de una norma contra el robo de caramelos, que corresponde a la expresión "estuvo mal robar los caramelos", y reconoce un estado de creencia de que Dios te castigará, que corresponde a la expresión "Dios te castigará". Pero no puede haber ningún estado mental que corresponda de manera similar a la frase completa, "Si estuvo mal robar el caramelo, entonces Dios te castigará", porque alguien que afirma sinceramente esa frase no necesita ni aceptar

la norma anti-robo de caramelos, ni creer que Dios te castigará. Los únicos recursos psicológicos de Gibbard son las creencias y las normas, pero no hay ninguna combinación posible de creencias y normas que pueda tomarse como correspondiente al condicional.

Dicho de otro modo: Gibbard quiere explicar el significado del condicional, "Si estuvo mal robar los caramelos, entonces Dios te castigará", por referencia al conjunto de mundos fáctico-normativos que el estado mental del hablante descarta. Como vimos anteriormente, el conjunto correcto de mundos a descartar es $\{(w_2, n_2)\}$. Pero no hay ninguna combinación posible de estados mentales -utilizando los recursos a disposición de Gibbard- que descarte sólo este mundo. Los posibles estados mentales relevantes, en la teoría de Gibbard, son:

Creer que Dios te castigará.

Creer que Dios no te castigará.

No tener opinión sobre si Dios te castigará.

Aprobar el robo de caramelos.

Desaprobar el robo de caramelos.

Ser neutral ante el robo de caramelos.

A partir de estos estados mentales se pueden formar nueve combinaciones coherentes. Ninguna de ellas tiene el resultado deseado de descartar sólo (w_2, n_2). Por ejemplo, supongamos que creo que Dios te castigará y apruebo el robo de caramelos. Esto no funciona, ya que descarta (w_2, n_1) De hecho, ninguna combinación que incluya creer que Dios te castigará funciona, ya que cualquier combinación de este tipo excluye (w_2, n_1). Tampoco funciona ninguna combinación que incluya creer que Dios no te castigará, ya que descartaría (w_1, n_1). Ninguna combinación que incluya aprobar el robo de caramelos funciona, ya que

descartará (w_1, n_2). Ninguna combinación que incluya desaprobar el robo de caramelos funciona, ya que descartará (w_2, n_1). Sólo queda la combinación de no tener opinión sobre si Dios te castigará y ser neutral respecto al robo de caramelos. Pero esto no descarta (w_2, n_2).

El problema aquí es la separación entre los estados de creencia y los estados de aceptación de normas, que es inevitable porque son diferentes tipos de estados mentales sin objetos comunes. En general, no es posible construir una actitud hacia una proposición compuesta a partir de actitudes hacia los componentes separados de la proposición. Por eso Gibbard no puede, en general, proporcionar significados para enunciados compuestos en los que un componente es evaluativo y el otro no evaluativo.

Nótese que el cognitivista no tiene ningún problema aquí. El cognitivista no necesita construir la creencia de que si robar el caramelo estuvo mal, entonces Dios te castigará a partir de opiniones separadas sobre el robo del caramelo y sobre el posible castigo futuro de Dios. En su lugar, el cognitivista tiene un único objeto, la proposición de que si robar el caramelo estuvo mal, entonces Dios te castigará, y un único estado mental (una creencia) que puede dirigirse hacia esa proposición.

*2.6 Soluciones de Blackburn

Una sugerencia debida a Simon Blackburn podría ayudar a Gibbard. Sin embargo, si lo hace, haría innecesaria la discusión de los mundos fáctico-normativos; por lo tanto, discutiré esto como una solución propuesta separada al problema Frege-Geach. La sugerencia de Blackburn es que al pronunciar un enunciado condicional, uno expresa un estado mental disposicional: una disposición a aceptar el consecuente, si uno fuera llevado a aceptar el antecedente. Así, si digo: "Si robar los

caramelos estuvo mal, entonces Dios te castigará", estoy expresando mi disposición a creer que Dios te castigará, si llego a desaprobar el robo.[28] Presumiblemente, Blackburn diría lo mismo de los condicionales que no contienen evaluaciones (la capacidad de hacerlo sería una virtud importante de esta teoría). Así, "Si mañana llueve, Snoopy se mojará" expresa una disposición a aceptar que Snoopy se mojará, si uno se convence de que mañana lloverá.

He aquí una historia para ilustrar por qué esto es erróneo. Un psiquiatra tiene un paciente que cree que (el paciente) está muerto. El psiquiatra le dice al paciente: "Muy bien. Usted sabe que la sangre de los muertos se coagula, ¿verdad?" Paciente: "Sí". "Entonces, si usted está muerto, si le pincho en el dedo, la sangre no saldrá", "Sí". El psiquiatra saca una aguja, pincha el dedo del paciente y la sangre sale. Psiquiatra: "¿Qué me dice ahora?" Paciente: "Bueno, creo que estaba equivocado. Supongo que la sangre de los muertos no se coagula".

En este ejemplo, el psiquiatra intenta involucrar al paciente con la siguiente inferencia:

> Si el paciente está muerto, su sangre se ha coagulado.
>
> La sangre del paciente no se ha coagulado.
>
> Por lo tanto, no está muerto.

El paciente acepta inicialmente la primera premisa. Se le lleva a aceptar la segunda premisa. Pero en lugar de aceptar la conclusión, simplemente revisa su creencia en la primera premisa.

Del mismo modo, si una persona acepta "Si *A*, entonces *B*", no tiene por qué estar dispuesta a aceptar *B* al convencerse de *A*; en lugar de ello, podría estar dispuesta, al convencerse de *A*, a dejar de creer "Si *A*, entonces *B*,[29] En algunos casos, esto sería perfectamente racional. Por

ejemplo, supongamos que yo creo que, puesto que Ana es un juez moral fiable, si Ana piensa que una persona es mala, entonces esa persona es mala. Pero cuando Ann declara inesperadamente que la Madre Teresa era malvada, no acepto que la Madre Teresa fuera malvada, sino que renuncio a la creencia de que si Ann dice que alguien es malvado, entonces lo es. Esto demuestra que no podemos interpretar que aceptar un condicional signifique estar dispuesto a aceptar el consecuente si uno llega a aceptar el antecedente.

Una salvedad que aparece en el enunciado de la teoría de Blackburn podría ser de ayuda en este caso:

> Supongamos que yo sostengo que, o bien Juan tiene la culpa, o bien no cometió el hecho. Entonces me encuentro en un estado en el que si una de las partes se me cierra, debo cambiar a la otra, o retirar el compromiso.[30]

La frase en cursiva parece dejar escapar a Blackburn de mi contraejemplo. Blackburn podría decir que cuando uno afirma sinceramente un condicional, está dispuesto, al aceptar el antecedente del condicional, a aceptar su consecuente o a retirar su compromiso con el condicional. En los dos ejemplos anteriores, ocurre lo segundo.

Recordemos que el objetivo original de la propuesta de Blackburn era permitir al no cognitivista explicar lo que significa aceptar un enunciado condicional con un antecedente evaluativo. Si bien puede ser cierto que una persona que acepta "Si *p* entonces *q*" está dispuesta, al convencerse de *p*, a aceptar *q* o a renunciar a "Si *p* entonces *q*", esto no puede tomarse como una explicación de lo que significa "Si *p* entonces *q*". El efecto de incluir la frase "o retirar el compromiso" dentro de la caracterización de aquello a lo que se supone que el compromiso se compromete, es vaciar de contenido el supuesto compromiso. Es como si yo dijera: "Por la presente prometo pagarte 10.000 dólares o retirar

esta promesa". Un compromiso de (realizar el acto *A* o retirar este compromiso) es equivalente a un "compromiso" de (realizar *A* o no), que no es ningún compromiso. Del mismo modo, una disposición a (aceptar *q* o dejar de tener esta disposición) es difícil de distinguir de una disposición a (aceptar *q* o no).

Para plantear el problema de otro modo, la afirmación de que una persona tiene alguna disposición debería implicar algunas predicciones sobre lo que la persona probablemente haría (en términos no disposicionales) al menos en algunas circunstancias; de lo contrario, tener la disposición sería indistinguible de no tenerla. Es decir, las afirmaciones sobre las disposiciones de una persona deben conllevar predicciones determinadas sobre las acciones, aparte de las "acciones" de adquirir y perder disposiciones. La disposición hipotética de Blackburn viola esta restricción. Una vez que se añade "o retirar el compromiso" a la caracterización de lo que uno está dispuesto a hacer, ya no hay ninguna predicción sobre en qué estados (no disposicionales) sería probable que una persona entrara en cualquier circunstancia. Ahora es perfectamente coherente con el análisis de Blackburn que una persona que tenga la supuesta disposición se comporte exactamente de la misma manera (caracterizada no disposicionalmente) en todas las circunstancias posibles que una persona que carezca de la disposición. La razón por la que la calificación de Blackburn le permite escapar de mi contraejemplo, entonces, es que vacía su análisis de contenido.

Mientras que la teoría inicial del significado de los condicionales puede parecer plausible a primera vista, el punto de vista de Blackburn sobre las disyunciones, expresado en la cita anterior, no lo es. Sostiene que aceptar una disyunción es estar dispuesto a aceptar cualquiera de las dos disyunciones si la otra te está vedada. Por ejemplo, acepto la afirmación "O hay unicornios viviendo en el planeta Júpiter, o las vacas

tienen cuatro estómagos". (Según Blackburn, esto significa que estoy dispuesto a aceptar que hay unicornios en Júpiter si descubro que las vacas no tienen cuatro estómagos. Pero no es ni remotamente así.

Blackburn propuso anteriormente una teoría ligeramente diferente, antes de llegar a la que hemos estado discutiendo.[31] Propuso que los enunciados compuestos con componentes evaluativos expresan una especie de evaluación de segundo orden; a grandes rasgos, expresan aprobación hacia el tipo de disposiciones que hemos estado discutiendo. Así, "Si robar el caramelo estuvo mal, Dios te castigará" expresa la aprobación del hablante hacia el hecho de tener (una disposición a aceptar que Dios te castigará, si uno llega a desaprobar el robo del caramelo). La diferencia entre los puntos de vista anterior y posterior de Blackburn es que en el primero dice que expresamos aprobación hacia ciertas disposiciones, mientras que en el segundo dice simplemente que expresamos las disposiciones mismas. ¿Afecta esta diferencia a la fuerza de mis objeciones? No; en esencia, se aplican las mismas dos objeciones. En primer lugar, aceptar un condicional no implica aprobar (una disposición a aceptar el consecuente si se acepta el antecedente). En su lugar, uno podría aprobar sólo una disposición a renunciar al condicional al adoptar el antecedente. En segundo lugar, el razonamiento de Blackburn no puede manejar disyunciones. "Robar está mal" implica lógicamente "Robar está mal, o los unicornios viven en Júpiter". Pero seguramente mi desaprobación del robo no me compromete a aprobar una disposición a aceptar que los unicornios viven en Júpiter, si uno llegara a aprobar el robo.

Incluso dejando de lado estos puntos, las propuestas de Blackburn hasta ahora podrían, como mucho, proporcionar una explicación de lo que sucede en frases como "Si el placer es bueno, entonces el helado de canela es bueno" y "O Simpson no mató a Brown, o el veredicto fue

injusto". Todavía no sabemos cómo entender, desde un punto de vista no cognitivista, oraciones como las siguientes:

1. Creo que el helado de canela es bueno.
2. Me pregunto si el aborto es malo.
3. Espero haber hecho lo correcto.
4. La gente no siempre hace lo correcto.
5. No recompenses a la gente por hacer cosas malas.

Analicemos cada una de estas cinco frases.

(1) "Creo que el helado de canela es bueno". ofrece Blackburn: "Creer que *X* es bueno o correcto es más o menos tener una valoración adecuadamente favorable de *X*".[32] Obsérvese que esto no es lo que normalmente significa "creer". Blackburn plantea una ambigüedad en la palabra "creer": cuando digo que creo que el placer es bueno, según Blackburn, en realidad no estoy diciendo que crea nada en el sentido en que diría que creo que va a llover mañana. Consideremos una forma de comprobar la ambigüedad; yo la llamo la prueba del chiste. Alguien dice: "Estefanía tiene la piel muy clara. Su piel es más clara que una pluma". La segunda frase suena graciosa porque trata dos significados de "claro", claro por oposición a oscuro y claro por oposición a pesado, como si fueran el mismo. Ahora contrasta el comentario, "Jerry es muy pro-vida. Cree que el aborto es un error, incluso más firmemente que que 2+2 son 4". Este comentario no tiene ninguna gracia. Pero si "creer" fuera ambiguo, de tal manera que uno no pudiera "creer" que el aborto es malo en el mismo sentido que uno puede "creer" que 2+2 son 4, entonces el segundo ejemplo sonaría gracioso de la misma manera que el primero.

(2) "Me pregunto si el aborto está mal". Blackburn dice: "Preguntarse si *X* es bueno o correcto es preguntarse qué doler, qué admirar o valorar".

Pero aunque me pregunto si el aborto está mal, afortunadamente no me pregunto si debo abortar. ¿Me pregunto si condenar el aborto? Tal vez, pero eso sería otra cuestión: preguntarse si condenar el aborto sería, en todo caso, preguntarse si condenar el aborto es correcto, no preguntarse si el aborto es correcto. La propuesta de Blackburn resulta aún menos plausible cuando se aplica a preguntas sobre el pasado o sobre otras personas: si "me pregunto si es correcto que yo haga *A*" significa "me pregunto si debo hacer *A*", entonces "me pregunto si fue correcto que Truman bombardeara Hiroshima" significa... ¿qué?

(3) "Espero haber hecho lo correcto". Blackburn no ofrece ninguna ayuda en este caso, y no veo ninguna interpretación de esta frase paralela a sus tratamientos de los casos anteriores, ni ninguna otra interpretación no cognitivista.

(4) "La gente no siempre hace lo correcto". Blackburn no discute este caso, pero siguiendo el ejemplo de sus tratamientos de (1) y (2), podríamos intentar esta paráfrasis: "La gente no siempre hace lo que hay que hacer".

Pero no tiene sentido. Podríamos decir que la frase expresa desaprobación hacia algunas cosas que la gente hace, pero es difícil ver a alguien expresando desaprobación sólo hacia "algunas cosas" pero no hacia nada en particular.

(5) "No recompenses a la gente por hacer cosas malas". Podríamos intentar interpretar esta frase como una expresión de desaprobación hacia la práctica de recompensar a la gente por hacer cosas que uno desaprueba. Pero esto sería tratar "cosa mala" como un término descriptivo equivalente a "cosa que uno desaprueba", lo cual es abandonar el no cognitivismo en favor del subjetivismo. La propuesta de que la frase expresa desaprobación hacia la práctica de recompensar a la

gente por hacer cosas que el hablante desaprueba, revierte de forma similar al subjetivismo.

En cambio, el punto de vista cognitivista es simple y directo. No requiere plantear ninguna ambigüedad en palabras como "creer", ni idear teorías semánticas excéntricas para cada una de las frases "problemáticas". Según el argumento cognitivista, todas estas frases son fácilmente comprensibles y significan justo lo que parecen significar.

*2.7 El no descriptivismo asertivo de Timmons

Desde sus inicios, el no cognitivismo, en sus esfuerzos por eludir las objeciones, se ha ido pareciendo cada vez más al realismo moral. Los no cognitivistas simples, como Ayer, se limitaban a decir que los enunciados morales no afirman nada, no expresan creencias y no pueden ser ni verdaderos ni falsos; tampoco existen propiedades o hechos morales. Los no cognitivistas posteriores, como Hare y Blackburn, intentaron calmar el sentido común permitiéndonos aplicar las palabras "verdadero" y "falso" a los enunciados morales y hablar de argumentos válidos que los incluyeran, pero manteniendo el resto de la doctrina, es decir, que los enunciados morales no afirman proposiciones ni expresan creencias, ni existen propiedades o hechos morales. Este truco filosófico requería cierta reinterpretación de "verdadero" y "falso", pero se consideraba aceptable. El objetivo del juego, para muchos filósofos, es idear interpretaciones de palabras y frases comunes en inglés que hagan que la mayoría de las afirmaciones que la gente corriente aceptaría se consideren verdaderas -o al menos apropiadas- sin dejar de mantener los puntos de vista metafísicos preferidos (en este caso, la opinión de que no existen valores objetivos). Una de las reglas del juego es no llamar nunca a las interpretaciones reinterpretaciones; siempre se dice que se

están dando los significados actuales de las palabras relevantes. El obispo Berkeley inauguró esta práctica, intentando demostrar que su negación de la existencia de la materia era perfectamente coherente con el sentido común,[33] y los filósofos contemporáneos la han elevado a la categoría de arte.

Volviendo a la historia del no cognitivismo, todavía no se consideraba que la teoría tuviera éxito, porque la gente corriente aceptará afirmaciones como: "Creo que comer vacas está mal", "Jon afirmó que comer vacas está mal", "El hecho de que la guerra sea injusta explica por qué la gente protesta contra ella" y "La esclavitud tiene la propiedad de la injusticia". El no cognitivismo, incluso de la variedad Hare-Blackburn, todavía tenía que considerar esas afirmaciones como falsas, si no disparatadas.

Mark Timmons ha dado recientemente el siguiente paso, y presumiblemente el último, en la asimilación no cognitivista de las formas realistas de hablar. Su "no descriptivismo asertivo" repudia -al menos verbalmente- casi todas las tesis que tradicionalmente definían el no cognitivismo. Permite que los enunciados morales sean afirmaciones, expresen creencias y puedan ser verdaderos o falsos. Hay propiedades y hechos morales. Timmons incluso repudia el relativismo: la verdad moral no depende de mis actitudes ni de las de los demás. Sin embargo -esta es la parte que hace de Timmons un antirrealista- las afirmaciones morales no son descriptivas, sino evaluativas. Las afirmaciones "descriptivas" son aquellas que pretenden describir estados de cosas en el mundo y, por tanto, requieren para su verdad que el mundo sea de una determinada manera; por el contrario, las "afirmaciones evaluativas" funcionan para guiar la acción y expresar un estado mental que consiste en "adoptar una postura moral sobre alguna cuestión".[34]

Si esta combinación de tesis le parece incoherente, es posible que

esté trabajando con malentendidos sobre el significado de algunas de las palabras del párrafo anterior, como "verdadero", "afirmación", "creencia", "propiedad" y "hecho". Empecemos por "verdadero". La visión tradicional de la verdad es la teoría de la correspondencia: para que una afirmación sea verdadera, debe corresponder a los hechos de un mundo exterior independiente de la mente. Por conveniencia, adoptemos la convención notacional de usar todas las mayúsculas cuando nos refiramos a los habitantes de este mundo independiente de la mente; así, OBJETOS, PROPIEDADES y HECHOS son el tipo de cosas que ocupan el MUNDO según el punto de vista de la correspondencia, y con las que nuestras afirmaciones tienen que corresponderse.[35]

Timmons, siguiendo algunas tendencias filosóficas recientes, prefiere una interpretación "minimalista" de la verdad. Para que un enunciado sea verdadero, dice, basta con que sea "correcto" afirmarlo, según cualesquiera que sean las normas para la clase de enunciados a la que pertenece. Para algunos tipos de afirmaciones, las normas que tenemos para saber cuándo son correctas incluyen que el MUNDO externo debe ser de una determinada manera. Pero las normas para otros tipos de enunciados -incluidos los evaluativos- no imponen tales exigencias al MUNDO. Cuando digo: "Es cierto que comer carne está mal", en realidad sólo estoy diciendo: "Comer carne está mal". Esta última afirmación no impone ninguna exigencia al MUNDO (de nuevo, sólo expresa mi postura moral), por lo que tampoco lo hace la primera.[36]

A continuación, "aserción": para Timmons, cualquier frase que tenga todos los marcadores lingüísticos de una aserción, del tipo tratado en la sección 2.3, es ipso facto una aserción; eso es todo lo que hay que hacer para ser una aserción. Así pues, los enunciados morales son asertivos. Del mismo modo, hay creencias morales, porque cualquier estado mental que pueda expresarse mediante una aserción cuenta

automáticamente como una "creencia". Hay hechos morales, porque todo lo que se requiere para que haya un "hecho" (pero no un HECHO) es que haya una aserción verdadera. Hay propiedades morales (no PROPIEDADES), porque todo lo que se requiere para que haya "propiedad" es que haya un adjetivo que apliquemos correctamente a las cosas. Así, Timmons consigue decir casi todas las cosas que suelen decir los realistas. Sus afirmaciones sólo difieren de las del realista moral cuando utiliza letras mayúsculas: entonces dice que no hay HECHOS ni PROPIEDADES morales, ni afirmaciones morales VERDADERAS.

Esto es sólo un resumen, pero basta para sentar las bases de mis críticas. Creo que la mejor manera de entender lo que ocurre en la obra de Timmons muestra también lo que hay de insatisfactorio en ella. En esencia, el punto de vista de Timmons viene a ser el mismo que el de los no cognitivistas tradicionales, pero con las siguientes estipulaciones sobre el uso de las palabras: Timmons amplía la palabra "aserción" para abarcar tanto lo que ellos llamaban "aserciones" como lo que ellos llamaban "expresiones emotivas"; amplía la palabra "creencia" para abarcar tanto lo que ellos llamaban "creencias" como lo que ellos llamaban "actitudes morales"; y amplía la palabra "verdadero" para abarcar tanto lo que ellos llamaban "creencias verdaderas" como lo que ellos llamaban "actitudes admirables".[37] Su opinión de que hay afirmaciones genuinas que no son descriptivas suena inicialmente radical, incluso paradójica; pero bien entendido, no está diciendo que los enunciados morales sean afirmaciones en el sentido en que los teóricos anteriores entendían la palabra "afirmación". Por supuesto, él diría que los pensadores anteriores entendieron mal la palabra inglesa "assertion", pero dejemos eso a un lado. Si el no descriptivismo asertivo sólo difiere verbalmente del no cognitivismo tradicional, entonces parece poco probable que ofrezca algún avance real en los problemas a los que se enfrenta el no cognitivismo.

De hecho, no lo hace. El tratamiento de Timmons del problema Frege-Geach es lo suficientemente parecido al de Blackburn, y su tratamiento de la verdad moral lo suficientemente parecido al de Hare, como para que no los discuta aquí por separado. En su lugar, plantearé dos objeciones específicas a sus puntos de vista.

La primera se refiere al error moral. A menos que seamos dogmáticos, todos admitimos que algunas de nuestras opiniones morales pueden estar equivocadas. Afirmaciones como "Creo que la clonación es buena, pero podría estar equivocado" son a menudo pertinentes. La interpretación realista de esa observación es sencilla, pero ¿cómo puede entenderla un antirrealista? ¿Cómo puedo equivocarme si no hay hechos morales independientes de mis actitudes? Timmons considera que éste es el problema más grave al que se enfrenta el antirrealismo. Su solución es que la admisión de la falibilidad moral expresa una actitud hacia las propias actitudes morales, a saber, "una actitud de apertura a la posibilidad de que [la propia perspectiva moral] pueda evolucionar hacia una perspectiva moral diferente como resultado de argumentos nuevos y persuasivos, así como de nuevas experiencias".[38]

Está claro que esto no es correcto. Lo que Timmons describe podría ser lo que se expresa con "puedo llegar a creer que mi perspectiva actual es errónea", pero no lo que se expresa con "mi perspectiva actual puede ser errónea". Consideremos una observación ligeramente diferente: "Mis puntos de vista morales pueden estar equivocados en formas que nunca descubriré".[39] Esto también tiene mucho sentido, pero evidentemente no expresa una actitud sobre la posibilidad de que yo cambie mis puntos de vista morales de alguna manera en el futuro. O supongamos que digo: "Puedo equivocarme al adoptar una perspectiva moral diferente como resultado de argumentos nuevos y persuasivos". Esto tiene sentido e implica que mi punto de vista moral puede evolucionar del modo que

describe Timmons, pero no implica que mi punto de vista actual pueda estar equivocado. Por lo tanto, uno puede admitir su falibilidad moral sin expresar su apertura a la posibilidad de que su punto de vista evolucione, y uno puede expresar su apertura a la posibilidad de que su punto de vista evolucione sin admitir su falibilidad moral.

Además, compare las dos frases:

> Creo que comer carne está mal, pero puedo estar equivocado.
>
> Creo que Napoleón fue derrotado en 1815, pero puedo estar equivocado.

La última frase, presumiblemente, sólo habla de cómo podría ser el mundo independiente de la mente, y no de posibles cambios futuros en mis creencias. Así que Timmons debe pensar que "podría estar equivocado" tiene significados diferentes en las dos frases. Si es así, entonces esto no tendría sentido:

> Así que hay dos cosas en las que podría estar equivocado: el estatus moral de comer carne y cuándo fue derrotado Napoleón.

Pero eso se deduce de las dos afirmaciones anteriores, y no carece de sentido.

Mi segunda objeción al argumento de Timmons es más fundamental. Timmons quiere demostrar que el sentido común no está comprometido con el realismo moral, ofreciendo interpretaciones de cómo funciona el lenguaje que nos permiten decir las cosas que normalmente decimos sin estar comprometidos con la existencia de HECHOS evaluativos. La estratagema fracasa porque sus interpretaciones simplemente no son lo que entendemos por "verdad", "hecho", etcétera. Propongo mostrar esto argumentando que, si fueran lo que queremos decir, entonces no podríamos entender la exposición de Timmons de su propia teoría; en

particular, no veríamos ninguna distinción entre el no-descriptivismo asertivo y el realismo moral.

Timmons sólo dispone de la lengua inglesa para explicar su teoría. Puede introducir nueva terminología técnica, pero luego debe explicar, en inglés, lo que significan sus términos técnicos. Ahora bien, la distinción que postula entre propiedades y PROPIEDADES es crucial para explicar su teoría y en qué se diferencia del realismo moral; a grandes rasgos, Timmons no es realista porque cree en las propiedades morales, pero no en las PROPIEDADES morales. ¿Cómo explica esta distinción? He aquí el pasaje completo que la introduce:

> *Según la teoría de la correspondencia, la verdad es una cuestión de correspondencia directa entre el lenguaje (o el pensamiento) y un mundo independiente de la mente de entidades, propiedades, relaciones o, más sencillamente, hechos (interpretados como constelaciones de los primeros). A efectos expositivos, será conveniente adoptar el método de Putnam de poner en mayúsculas palabras como "mundo", "objeto", "propiedad", y "hecho" para dejar claro que cuando hablo de los habitantes de la mente y del mundo independiente del discurso es parte del punto de vista de la correspondencia. Como veremos más adelante, el irrealista moral puede ayudarse a sí mismo para hablar de propiedades, relaciones y hechos morales, pero, por supuesto, es parte de la historia del irrealista que no hay PROPIEDADES, RELACIONES o HECHOS morales. Así pues, según el punto de vista de la correspondencia, la verdad de cualquier oración es una cuestión de su correspondencia con el MUNDO de los OBJETOS, PROPIEDADES, RELACIONES y (si se quiere hablar así) HECHOS.*[40]

Nótese que todo lo que se nos dice sobre la naturaleza de las PROPIEDADES es que son propiedades independientes de la mente y del lenguaje.

En lo sucesivo, utilizaré el término "Timmons-property" para referirme a lo que Mark Timmons entiende por la palabra "property" cuando la pone en minúscula (es una pregunta abierta, hasta ahora, si eso es lo mismo que lo que significa "property" en inglés). Ahora bien, yo sostengo que las opiniones de Timmons no tienen sentido, a menos que esté utilizando mal la palabra "propiedad". Consideremos el significado de "property" en inglés estándar: ¿significa PROPIEDAD o significa propiedad de Timmons?

Caso 1: Supongamos que "property" en inglés significa PROPIEDAD. Entonces el pasaje anterior de Timmons da una definición circular: en efecto, nos dice que una PROPIEDAD es una PROPIEDAD independiente de la mente. Y lo que es más importante, la afirmación central de Timmons de que hay propiedades morales pero no PROPIEDADES morales es una contradicción directa.

Se podría decir que no se contradice, porque cuando dice esto, está utilizando "propiedades" de manera no estándar (es decir, para referirse a las propiedades de Timmons). Pero en ese caso, su explicación del significado de las palabras es incorrecta y, por tanto, fracasa en su intento de demostrar que el sentido común no está comprometido con el realismo moral.

Caso 2: Supongamos que "propiedad" en inglés significa Timmons-property. Entonces las PROPIEDADES, según el pasaje anterior, comprenden el subconjunto de propiedades de Timmons que son independientes de la mente. Timmons no define "independiente de la mente", pero aquí hay dos criterios razonables:

a) La *F*-sidad es independiente de la mente, siempre que: Si una cosa es *F*, seguiría siendo *F* aunque tuviéramos actitudes completamente diferentes sobre ella.[41]

b) La *F*-sidad es independiente de la mente, siempre que: Es posible que una cosa sea F aunque nadie piense así.

Timmons afirmaría la independencia mental de las propiedades morales en ambas interpretaciones. Afirma que tenemos derecho a admitir la falibilidad moral, incluso de los tipos más radicales, por lo que aceptaría que las propiedades morales satisfacen (b). También subraya que su punto de vista no es relativista y que no sostiene que una afirmación moral describa las actitudes del hablante. Piensa que las afirmaciones morales son "categóricas", lo que implica, entre otras cosas, que el hablante debe ser capaz de apoyarlas con "razones no subjetivas, en otras palabras, razones que no apelen a cuestiones de gusto individual, preferencia personal o similares".[42] Así que parece que, desde su punto de vista, la lógica del discurso moral obliga a afirmar que las propiedades morales satisfacen (a). Además, esto es obvio en cualquier caso: es falso que "la esclavitud no estaría mal, si nosotros, los que odiamos la esclavitud, nos tranquilizáramos un poco".[43] Así que parece que, si una PROPIEDAD es sólo una propiedad de Timmons que no depende de la mente, entonces Timmons sí considera que hay PROPIEDADES morales. Por tanto, es un realista moral.

Ahora bien, esto es lo que creo que está pasando realmente. Por supuesto que Timmons no es realista. Entendemos lo que dice Timmons, y entendemos que difiere del realismo, porque "property" en inglés significa PROPIEDAD; las mayúsculas no añaden nada. Entendemos que el punto de vista real de Timmons es que (a) en realidad no hay propiedades morales, pero (b) podemos seguir usando la palabra "propiedad", porque una de las reglas del lenguaje moral es que se puede usar esa palabra aunque no haya propiedades a las que se refiera. La razón por la que el punto de vista de Timmons es inicialmente difícil de entender y por la que aparece tan tarde en escena, es que las reglas del

discurso filosófico le impiden describir su punto de vista de la manera más perspicua. Porque si describe su punto de vista como yo lo he hecho, su falsedad es evidente: no es el caso que en inglés se permita decir que hay propiedades de algún tipo cuando en realidad no las hay.

Si los puntos de vista semánticos de Timmons fueran correctos, entonces no podríamos entender qué es el realismo. Sus puntos de vista le dan derecho -de hecho, le exigen- a afirmar todas las frases características del realismo: "Hay hechos morales", "Los hechos morales son independientes de nuestras actitudes", "Los enunciados morales son verdaderos cuando se corresponden con los hechos",[44] etc. Si los términos de las letras mayúsculas no fueran los términos ordinarios del inglés -si fueran términos técnicos, como Timmons debe representar que son- entonces tendría que explicar en inglés lo que significan. Pero ninguna de sus explicaciones tendría éxito, porque interpretaríamos todas las expresiones de esas explicaciones de forma minimalista, por lo que consideraríamos que el realista no se compromete a más de lo que se compromete Timmons. Cuando Timmons intentara decirnos que los realistas creen en un MUNDO INDEPENDIENTE DE LA MENTE, simplemente pensaríamos que se refiere a un mundo independiente de la mente de Timmons.

Las versiones recientes del no cognitivismo son una especie de engaño metalingüístico. Pretenden abrazar el sentido común en las palabras, pero rechazarlo en el fondo. Nadie excepto Berkeley se dejó engañar cuando Berkeley afirmó ser un anti-escéptico; tampoco deberían engañarnos los no-cognitivistas de hoy en día.

2.8 La evidencia introspectiva

He caracterizado el cognitivismo y el no cognitivismo como tesis sobre

el significado de los enunciados evaluativos. Asumiendo que los enunciados evaluativos expresan juicios de valor, las afirmaciones semánticas también implican afirmaciones psicológicas: El cognitivismo implica que lo que llamamos juicios de valor son una especie de creencia, como sugiere el término "juicio"; el no cognitivismo implica que, en cambio, son una especie de emoción, deseo u otro estado no cognitivo. Podríamos esperar resolver esta disputa mediante la introspección. Cuando hacemos juicios de valor, ¿la experiencia se parece más a las típicas emociones, deseos u otras actitudes no cognitivas, o se parece más a las típicas experiencias de formación de creencias?

El emotivista podría citar el hecho de que normalmente experimentamos emociones cuando emitimos juicios morales. De hecho, las disputas sobre cuestiones como la moralidad del aborto y la justicia de la pena capital se encuentran entre las más emocionales de todas las disputas. Esta es una prueba a favor del no cognitivismo; después de todo, si en general no hubiera emociones asociadas a los juicios morales, sin duda lo consideraría una prueba a favor del cognitivismo. Pero antes de apoyar el no cognitivismo por este motivo, busquemos una visión más matizada de lo que las dos teorías predicen sobre nuestra psicología.

En primer lugar, aunque el no cognitivismo predice más claramente que el cognitivismo una asociación entre el juicio moral y la emoción, esta asociación no es inexplicable desde un punto de vista cognitivista. Sólo demuestra que nos preocupan las cuestiones morales, en algunos casos con mucha intensidad. Hay muchas otras cuestiones que también nos importan mucho. Los debates sobre si Dios existe, quién disparó a JFK o qué significa una cláusula concreta de un contrato pueden suscitar emociones fuertes. Pero esto ejerce poca o ninguna presión a la hora de considerar las afirmaciones sobre esos temas como meras expresiones

emocionales desprovistas de contenido fáctico.

En segundo lugar, nuestras emociones sobre algunas cosas son más fuertes que sobre otras. Si la principal explicación de la correlación entre el juicio moral y la emoción es que "bueno" y "malo" son sólo términos que expresan emociones, entonces ¿qué deberíamos predecir sobre cómo debería variar la fuerza de las emociones morales? Parece que la fuerza de las emociones asociadas a los juicios morales debería ser aproximadamente proporcional al nivel percibido de bondad o maldad de una cosa y a la confianza en el propio juicio moral: cuanto peor se considere una cosa mala y más confianza se tenga en el propio juicio, más fuerte debería ser el sentimiento al respecto. ¿Es esto cierto? En algunos casos, sí: creo más firmemente que el asesinato es malo que cruzar la calle imprudentemente. En otros casos, no: siento algo más fuerte cuando alguien me roba el equipo de música (y ésta es una emoción moral, la indignación) que cuando el emperador Nerón ejecuta a Octavia en el año 62 d.C. A decir verdad, siento muy poco o nada por el asesinato de Octavia. Pero la ejecución de Octavia por Nerón fue mucho más mala que el robo de mi equipo de música, y estoy bastante seguro de ello. Esto sugiere que las emociones a menudo asociadas con las creencias morales tienen más que ver con cosas como nuestra preocupación por nuestros propios intereses, los grupos sociales con los que nos gusta asociarnos y las imágenes de nosotros mismos que nos gusta mantener, que con un significado emotivo incorporado al término "malo". Aunque no puedo catalogar todas las cosas que influyen en nuestras emociones, no está claro por qué el tipo de factores que explican por qué la gente se emociona ante otros asuntos no pueden explicar también nuestras emociones ante las cuestiones morales.

En tercer lugar, podemos distinguir entre el grado de confianza de un juicio moral y el grado de acierto o error que atribuye el juicio moral.

Imaginemos a alguien que piensa que el aborto es probablemente incorrecto, pero que no confía en ese juicio. Ve que hay argumentos plausibles en el otro lado. Sin embargo, dice, si el aborto está mal, está muy mal, ya que la única forma de que estuviera mal sería que fuera similar al asesinato. En este caso, el grado de maldad percibido es muy alto, pero la confianza en el juicio es muy baja. Según la teoría emotivista, ¿a qué podrían corresponder estas dos dimensiones? En la teoría emotivista, uno simplemente experimenta un sentimiento sobre el aborto. Este sentimiento puede tener distintos grados de intensidad: la gente puede sentir algo con más o menos fuerza. Pero ¿cómo podría esto dar lugar a nuestra sensación de que aquí hay dos cosas que vienen en grados y que pueden tener grados radicalmente diferentes en un caso dado?

En cuarto lugar, muchos juicios morales se hacen desapasionadamente, sin emoción evidente. Ahora bien, en el caso del asesinato de Octavia por Nerón, el emotivista podría decir que tengo un sentimiento general de desaprobación hacia el asesinato, y esto me induce a decir que el asesinato de Octavia estuvo mal aunque no tenga ningún sentimiento hacia Octavia en particular. Pero esto nos lleva a una característica desconcertante (para los emotivistas) de las emociones morales: los principios morales más fundamentales tienden a provocar menos sentimientos. Consideremos los siguientes candidatos a principios morales fundamentales:

> Disfrutar es bueno.
>
> Uno debe tratar a los demás como desearía ser tratado.
>
> La acción correcta es siempre la que produce más bien.
>
> Un ser sólo tiene derecho a algo si puede desearlo.

Sean o no correctos estos principios, son ejemplos del tipo de cosas que

los filósofos consideran, de forma bastante plausible, principios morales fundamentales. Pero poca gente se emociona con este tipo de afirmaciones. Tienden a emocionarse mucho más ante ciertos juicios morales derivados: que Steve maltrató a Alicia, que el veredicto de los Simpson fue injusto o que las exigencias del jefe no son razonables. Esto no sorprende al cognitivista, ya que no es más que un ejemplo de un patrón general: la gente se muestra más emocional ante asuntos concretos que afectan a sus propios intereses y a los de las personas que conoce que ante cuestiones abstractas y teóricas. Pero para el emotivista el fenómeno es inesperado: si los juicios morales son esencialmente emocionales, ¿no deberían los juicios morales más fundamentales mostrar las emociones morales más claras y puras?

Nótese aquí una asimetría importante. El cognitivismo no exige que los juicios morales sean siempre desapasionados; sabemos que las personas pueden verse influidas por las emociones a la hora de formar sus creencias, del mismo modo que sus creencias pueden a su vez influir en sus emociones; no hay nada especial en la moral a este respecto. Pero el emotivismo exige que los juicios morales sean siempre emocionales, porque, según la teoría, eso es precisamente lo que son los "juicios" morales. La introspección demuestra que esto es falso.

En quinto lugar, los procedimientos por los que, al menos a veces, formamos juicios morales -sobre todo, en filosofía moral- se parecen mucho más a la forma en que formamos creencias que a la forma en que solemos adquirir emociones. Quiero saber si el aborto es realmente malo. Acudo a la literatura de filosofía moral sobre el tema. Leo el experimento mental de Thomson sobre el violinista. Leo el argumento de Marquis sobre los derechos de las personas potenciales. Leo el argumento de Tooley sobre los criterios para tener derecho a la vida.[45] En todo esto, mi actividad mental es indistinguible de lo que sería

normalmente cuando estoy tratando de entender algo. Comienzo en un estado de suspensión del juicio. Sopeso las razones a favor y en contra de una conclusión. Por supuesto, nada impide intrínsecamente que este tipo de actividad provoque alguna emoción. Pero normalmente no es así como funcionan las emociones. Normalmente no acudo a la literatura profesional con el propósito de tener una emoción sobre algo. Normalmente no sopeso cuidadosamente los argumentos, buscando errores lógicos, con el fin de formarme una emoción. Normalmente no intento decidir qué emociones tener, y mucho menos me preocupo por "hacerlo bien". Ninguna de esas cosas es característica de cómo llego a tener miedo a las alturas, o a amar a mi novia, o a que me moleste Barney el dinosaurio.

Como en el caso de las pruebas lingüísticas, aquí todo parece ser exactamente como sería si una interpretación cognitiva del lenguaje moral fuera correcta. Concluyo que los términos éticos son cognitivamente significativos, como lo son "azul" y "cuadrado", y que los enunciados éticos son afirmaciones sobre un aspecto de la realidad, como lo son "El cielo es azul" y "Los perros son cuadrados".

Notas

1. Este no es el único punto de vista de la función de "bueno" consistente con el cognitivismo. Los cognitivistas con simpatías nominalistas podrían rechazar el relato, pensando que ninguna palabra se refiere a una propiedad. Incluso los cognitivistas que aceptan que algunos términos denotan propiedades podrían negar que "bueno" lo haga, siguiendo la escasa visión de las propiedades de Armstrong (1978b). Cualquier tipo de cognitivista mantendría, sin embargo, que "El placer es bueno" es asertivo del mismo modo que "El cielo es rojo".

2. Ayer (1952, capítulo 6).

3. Véase Hare (1952) para una declaración clásica del prescriptivismo, aunque Hare

afirma que los enunciados morales pueden ser verdaderos o falsos. Véase la discusión más adelante, sección 2.4.

4. Hare (1952).

5. Stevenson (1963, pp. 16, 64-5) atribuye a los enunciados morales las funciones tanto de expresar las actitudes del hablante como de influir en las actitudes de la audiencia.

6. Gibbard (1990, capítulo 4) postula un estado sui generis, no cognitivo, de "aceptación de la norma", que no es ni un tipo de creencia ni un tipo de emoción.

7. Pero véase más adelante, sección 2.7: Timmons podría estar en desacuerdo con esta caracterización.

8. Ayer (1952, p. 106-8).

9. Como admite Blackburn, "nadie niega que los fenómenos superficiales del lenguaje -el hecho de que usemos predicados morales y apliquemos la verdad o la falsedad a los juicios que hacemos cuando los usamos- suponen un problema para el proyectivismo. Por eso tientan al realismo" (1984, p. 196).

10. Los puntos (a)-(e) se discuten en Glassen (1959). Jackson, Oppy y Smith (2004) sostienen que el tipo de criterios sintácticos empleados en la presente sección son insuficientes para demostrar que una oración es apta para la verdad, esencialmente porque la oración debe tener también un patrón de uso suficientemente rico y debe utilizarse para expresar creencias. No afirmo aquí que cualquier combinación de los puntos (a)-(g) por sí sola implique que los enunciados éticos sean cognitivos. Más bien, afirmo que (a)-(g) hacen que las explicaciones no cognitivistas (como el emotivismo y el prescriptivismo) de los significados de los enunciados evaluativos sean altamente inverosímiles, y que una explicación cognitivista es la única alternativa plausible que queda. No es plausible, por ejemplo, que los enunciados éticos carezcan de un patrón de uso suficientemente rico para ser aptos para la verdad, o que no expresen ni creencias ni actitudes no cognitivas.

11. Las posibles excepciones son los llamados "performativos" (Austin 1975, pp. 4-6), como "Por la presente le condeno a muerte por electrocución" (dicho por un juez en una sala de tribunal), pero nótese que estos enunciados expresan, no obstante, proposiciones, aunque sirven para hacer que esas proposiciones sean verdaderas en lugar de afirmar que son verdaderas (Searle 1979, p. 16).

12. El problema se elabora en Geach (1965), que lo deriva de un punto de Frege. Searle (1969, pp. 136-41) proporciona una declaración especialmente clara.

13. Me centro en lo que puede aparecer como antecedente de un condicional. Las cláusulas no cognitivas pueden aparecer como consecuentes, como en "Si ves a Barney, golpéalo". (Compárese con Brighouse [1990, pp. 230-1].) Esto se debe a que uno puede emitir una orden condicional, una orden que tiene efecto condicionado a la obtención de algún hecho. No se puede, sin embargo, tener un hecho que sea de manera similar condicional a una orden, que es la razón por la que un imperativo no puede aparecer en el antecedente.

14. Compárese con Geach (1965, pp. 463-4).

15. El siguiente teorema de la probabilidad subraya esto: Si $P(e|h) > P(e|\sim h)$, entonces $P(h|\sim e) < P(h)$. Introduzca "no cognitivismo" para *h* y "Encontramos las frases enumeradas en (4) extrañas, malformadas o confusas" para *e*.

16. Hare (1963, pp. 21, 27-8; 1999, pp. 24-5).

17. Esto es una simplificación de la visión de Hare sobre el significado prescriptivo, ya que 1 no ha mencionado las restricciones de universalidad o consistencia de los imperativos. Pero esto no es importante para mi preocupación actual.

18. Para asegurar la relevancia del punto de vista de Hare, supongamos que la presencia de armas de destrucción masiva se considera parte de la base de superveniencia para la rectitud de la invasión. Naturalmente, toda la base de superveniencia, si George tiene puntos de vista éticos remotamente plausibles, tendría que ser un estado de cosas mucho más complejo; esto sólo sirve para hacer aún menos plausible la tesis de Hare de que George está describiendo la base de superveniencia.

19. Pero véase más adelante la otra teoría de la verdad moral de Hare (la teoría de la aprobación), que permitiría a Jean-Pierre calificar de "falsa" la afirmación de George.

20. Hare (1999, p. 25).

21. Hare (1999, p. 18); el énfasis es de Hare.

22. El último ejemplo es de Geach, quien refuta la teoría de la verdad del endoso en el mismo aliento que la teoría no cognitivista de la ética (1965, pp. 457,462-5).

23. Compárese con lo anterior, pp. 22-3. Nótese también que tal teoría subjetivista de la verdad sería singularmente inverosímil, implicando, entre otras cosas, que "Todo lo que apruebo es automáticamente verdadero" es una tautología, en lugar de ser la expresión del dogmatismo egoísta que parece ser. Véanse también los argumentos contra el subjetivismo ético en el capítulo 3.

24. Hare (1999, p. 18); el énfasis es de Hare.

25. Gibbard (1990, pp. 94-9). Gibbard formula en realidad su teoría como una teoría de la racionalidad. Un conjunto completo de normas sería aquel que emitiera un veredicto (racionalmente requerido, racionalmente opcional o irracional) sobre cada posible acción, creencia, sentimiento u otro objeto de evaluación racional. Técnicamente, en opinión de Gibbard, las normas no especifican directamente si una acción es moralmente correcta o incorrecta; especifican si es racional tener ciertas emociones morales hacia una acción. Sin embargo, esta distinción no es relevante para la cuestión de si Gibbard puede escapar al problema de Frege-Geach, y seguiré hablando como si las normas especificaran directamente lo que es moralmente correcto o incorrecto.

26. Lewis (1986, pp. 40-50). Hay objeciones a esta visión del significado, especialmente el problema de los enunciados lógicamente equivalentes que tienen significados diferentes, como "2+2=4" y "Todas las solteronas son solteras". Aunque este último problema puede refutar la explicación del significado de los mundos pósibles, supondré, en aras de la argumentación, que la explicación es aceptable. Aquí sólo quiero discutir los problemas del punto de vista de Gibbard que tienen algo especial que ver con la moralidad.

27. Aquí leo el condicional materialmente, por lo que ($P \rightarrow q$) " -(p & ~q). Si se hace una lectura más fuerte del condicional, creo que los problemas para Gibbard son aún mayores, pero bastará en esta sección con mostrar que Gibbard no puede interpretar ni siquiera un condicional material de este tipo.

28. Blackburn (1998, p. 71). Hare (1970, pp. 16-19) hizo un análisis similar, sosteniendo que el significado de la conectiva "si-entonces" se agota en sus propiedades lógicas. La propuesta de Blackburn es inicialmente más satisfactoria, ya que Blackburn nos dice qué estado mental es expresado por un condicional que contiene un componente evaluativo. En mi opinión, el ejemplo de Dreier (1996, pp. 42-3) que implica el acto de habla de 'abordar' demuestra la insuficiencia de la propuesta de Hare.

29. Van Roojen (1996, pp. 333-4) hace un punto similar, sobre las disyunciones.

30. Blackburn (1998, p. 71); segundo énfasis añadido.

31. Blackburn (1984, pp. 191-6).

32. Esta y la siguiente cita sobre el asombro son de Blackburn (1998, p. 70).

33. Berkeley ([1710]1965, secciones 33-40).

34. Timmons (1999, capítulo 4). Véanse especialmente las páginas 143, 153, 129, 138 y 149-50. Timmons (pp. 143-4) repudia la etiqueta de "no-cognitivista", pero esto es un punto meramente verbal; claramente pertenece a los otros discutidos en este capítulo.

35. Timmons (p. 116) recoge esta convención de Putnam. Aunque Timmons habla en términos de hechos mentales, sería más preciso hablar en términos de hechos que son independientes de su representación.

36. Compárese con la teoría de la verdad de Hare, discutida en la sección 2.4.

37. Dreier (2002, p. 165) lo ha señalado.

38. Timmons (1999, pp. 71, 167).

39. O, para anticipar posibles respuestas por parte de Timmons, considere "Algunos valores fundamentales míos podrían estar equivocados" y "Mi perspectiva moral podría estar equivocada en formas que no reconocería incluso después de una profunda reflexión". Ambas son claramente inteligibles.

40. Timmons (1999, p. 116).

41. Propongo esto como una condición suficiente, pero no necesaria, para la independencia de la mente. Supongamos que está mal quemar a tu vecino de al lado en efigie en tu patio delantero, porque esto heriría sus sentimientos. Es plausible que si su vecino tuviera una actitud totalmente diferente hacia la acción -por ejemplo, si deseara firmemente ser quemado en efigie y lo considerara un honor- entonces la acción ya no sería incorrecta.

Este examen, considero, no muestra que la ilicitud dependa de la mente en el sentido pretendido, mostrando que la condición (a) del texto no es necesaria para la independencia de la mente. Pero mientras haya al menos un caso de una acción incorrecta que sería incorrecta incluso si tuviéramos actitudes totalmente diferentes sobre ella en los aspectos relevantes, la incorrección es independiente de la mente. Véase la sección 1.2 para una caracterización más completa de la objetividad, que considero esencialmente equivalente a la independencia mental.

42. Timmons (1999, pp. 145-6). En las págs. 149-50, reniega del relativismo.

43. La cita es de Dreier (2002, p. 166), quien insiste en el punto actual de que las descripciones de Timmons de su punto de vista amenazan con colapsar en el realismo moral.

44. Wright, aunque es un minimalista de la verdad, enumera "ser verdadero es corresponder a los hechos" como una perogrullada parcialmente definitiva de "verdadero" (1992, p. 34). Timmons (1999, p. 153) está de acuerdo.

45. Thomson (1971); Marquis (1989); Tooley (1972).

3

Subjetivismo

3.1 ¿Qué es el subjetivismo?

Los subjetivistas consideran que las propiedades morales son reducibles a propiedades y relaciones psicológicas. Piensan que para que un objeto sea bueno es necesario que toda persona o grupo tenga (o esté dispuesta a tener) alguna actitud o reacción psicológica hacia él. He aquí algunos ejemplos de teorías subjetivistas; en cada caso, el signo "=" indica que la expresión de la derecha explica lo que significa que la expresión de la izquierda sea verdadera:[1]

a) *x* es bueno = El hablante cree que *x* es bueno.

b) *x* es bueno = El hablante aprueba *x*.

c) *x* es correcto = La sociedad aprueba *x*.

d) *x* es correcto = *x* está de acuerdo con los deseos de Dios.[2]

Las teorías (a) y (b) relativizan la moridad al individuo; es decir, suponen que dos individuos diferentes pueden llamar correctamente "buenas" a cosas diferentes. Según (a), cuando yo digo que algo es bueno, quiero decir que lo creo; cuando tú dices que algo es bueno, quieres decir que lo crees. La teoría (c) hace que la moralidad sea

relativa a una cultura; implica que los miembros de diferentes sociedades pueden llamar "buenas" a cosas diferentes, aunque dos miembros de una misma sociedad no.

Obsérvese que cada una de estas teorías es una teoría sobre lo que es ser bueno o correcto, no sobre lo que es ser considerado bueno o correcto. Por lo tanto, (b) no sostiene que para que una cosa sea tomada como buena es para que sea aprobada. Sostiene que para que una cosa sea buena es para que sea aprobada.

No hay que confundir el subjetivismo con la idea de que no hay verdades o hechos morales. Tomemos la teoría (c): es definitivamente cierto que nuestra sociedad desaprueba el robo; por tanto, un relativista cultural debería decir que es definitivamente cierto que robar es malo (en nuestra sociedad). Lo que el relativista niega es que esta verdad sea objetiva, es decir, independiente de los observadores en el sentido definido en la sección 1.2. Pero como ciertamente hay hechos sobre los observadores, no hay dificultad en reconocer los hechos morales.

El subjetivismo es una forma de reduccionismo sobre el valor. Los argumentos que despliego en el siguiente capítulo sirven para refutar el reduccionismo ético en general. Sin embargo, el subjetivismo es un punto de vista tan común que vale la pena dedicarle un capítulo adicional. Hay problemas especiales para el subjetivismo que no se aplican a otras versiones del reduccionismo. Analicemos cada una de las cuatro versiones del subjetivismo mencionadas anteriormente.

3.2 Subjetivismo individualista

Comienza con la teoría (a):

a) *x* es bueno = El hablante cree que *x* es bueno.

Su principal problema es de circularidad: para saber qué es que una cosa se crea que es *F*, primero debemos saber qué es *F*. Un problema relacionado es la regresión infinita: si [*x* es bueno] = [creo que *x* es bueno], entonces [creo que *x* es bueno] = [creo (creo que *x* es bueno)], y así sucesivamente. En general, es incoherente postular una proposición cuya verdad consista únicamente en que tú la creas. El mismo punto se aplica a la opinión de que ser bueno es ser creído como bueno por la sociedad, o por Dios, o por cualquier otra persona, y lo mismo se aplica si sustituimos "percibido", "conocido", "afirmado" o cualquier otro verbo por "creído". La palabra "bueno" no debe aparecer dentro de la explicación de lo que es que *x* sea bueno.

Conviene tener en cuenta este punto cuando lleguemos a la teoría (b):

b) *x* es bueno = El hablante aprueba *x*.

Si -como parece plausible- "aprobar" una cosa es tener una creencia moral sobre ella, como la creencia de que la cosa es buena, entonces (b) se derrumba en la teoría benigna (a) con toda su incoherencia concomitante. Por lo tanto, el subjetivista tendrá que decir que la "aprobación" denota algún tipo de actitud o emoción no analizable en términos de creencias morales. Aun así, la teoría (b) se enfrenta a graves problemas.[3]

En primer lugar, imaginemos que nos encontramos con un neonazi. Él dice:

1. Apruebo el asesinato de judíos.
2. Matar judíos es bueno.

Su primera afirmación es cierta. Pero la segunda obviamente no lo es. Pero la teoría (b) implica que ambas afirmaciones significan lo mismo,

por lo que ambas son verdaderas; de hecho, que estamos comprometidos por nuestra admisión de que (1) fue dicha con verdad a admitir que (2) fue igualmente dicha con verdad. Esto parece absurdo.

Por supuesto, (b) no nos compromete a decir "Matar a los judíos es bueno"; puesto que desaprobamos esa matanza, podemos decir con verdad que es mala. Pero al hacerlo, no contradecimos lo que dijo el neonazi. Informamos sinceramente de que matar a los judíos es malo para nosotros, pero sin embargo es bueno para los nazis. La teoría (b) hace, pues, imposible estar en desacuerdo con nadie sobre la moralidad: ésta es una segunda consecuencia absurda. Al decir que matar a los judíos es malo, obviamente pretendemos contradecir, y estamos contradiciendo, la afirmación del neonazi.

La tercera consecuencia absurda de la teoría (b) es que mis actitudes morales son infalibles.[4] Es decir, sería correcto afirmar que "todo lo que apruebo es automáticamente bueno". Sólo un egoísta dogmático pensaría esto, pero la teoría (b) implica que esto es necesariamente cierto, ya que el hecho de que algo sea bueno consiste simplemente en que yo lo apruebe. Si es así, puedo averiguar lo que es bueno y lo que es malo simplemente introspeccionando mis propias actitudes. Por tanto, expresiones como la siguiente serían una marca de confusión: "Lo apruebo, pero ¿es realmente bueno?". Sin embargo, esto no parece ser una observación confusa; al igual que puedo saber que creo en algo pero todavía tengo algunas dudas sobre si es verdad, puedo saber que apruebo algo pero todavía tengo algunas dudas sobre si es bueno.

En cuarto lugar, ¿por qué apruebo las cosas que apruebo? Si hay alguna razón por la que apruebo las cosas, entonces parece que esa razón, y no el mero hecho de mi aprobación, explica por qué son buenas. Si apruebo *x* por alguna característica que la hace deseable, admirable, etc., en algún aspecto, entonces la deseabilidad de *x* (etc.) sería un hecho

evaluativo previo a mi aprobación. Por otro lado, si apruebo *x* sin ninguna razón, o por alguna razón que no muestra que *x* sea deseable (etc.) en ningún aspecto, entonces mi aprobación es meramente arbitraria. ¿Y por qué el hecho de que alguien apruebe arbitrariamente algo lo convierte en bueno?

Estas razones pueden explicar por qué pocos filósofos han adoptado la teoría (b). Pasemos a la teoría más popular (c).

3.3 El relativismo cultural

Las teorías como (c) se denominan "relativismo cultural":

c) *x* es correcto = La sociedad aprueba *x*.

El relativismo cultural se enfrenta esencialmente a los mismos problemas que el subjetivismo individualista, con pequeñas modificaciones.

En primer lugar, imagine que usted es un ciudadano alemán durante la Segunda Guerra Mundial. Un oficial nazi afirma: "Matar a los judíos es correcto", y la mayor parte de su sociedad y de la tuya está de acuerdo.[5] Tú, sin embargo, eres un resistente: como Oscar Schindler, te dedicas en secreto a intentar evitar que se mate al mayor número posible de judíos. Según la teoría (c), el oficial nazi tiene razón y usted está equivocado, ya que él y no usted está actuando de acuerdo con las normas sociales imperantes. Lejos de ser un héroe moral, el relativista cultural debe decir que Schindler era inmoral. Es difícil ver la plausibilidad de este punto de vista.

Por supuesto, el relativista puede decir que Schindler fue un héroe según nuestras normas morales. Pero gran parte de la razón de ser de un

relativista cultural -quizás la motivación central del relativismo- ha sido tradicionalmente que uno se opone a juzgar otras culturas según nuestros estándares morales, o incluso a juzgar otras culturas. Por tanto, el hecho de que la Solución Final fuera mala según nuestros estándares era irrelevante para la gente que vivía en la sociedad nazi -no les daba ninguna razón para actuar de una manera particular- y el relativista cultural no tiene ninguna crítica que hacer a los estándares morales que permitieron a los nazis hacer lo que hicieron. De hecho, si un día las convenciones de nuestra propia sociedad cambian de tal manera que aprueben el asesinato en masa de algún grupo minoritario, el relativista cultural sólo puede aconsejarnos que, ese día, será moralmente correcto participar en el derramamiento de sangre.

En segundo lugar, ¿cómo explica la teoría (c) los desacuerdos morales? Si una persona dice "El aborto está mal" y otro miembro de la misma sociedad dice "El aborto no está mal", entonces están en desacuerdo. Pero, según (c), sólo están en desacuerdo sobre cuáles son las convenciones de su sociedad: están discutiendo si el aborto está de hecho permitido en su sociedad. Es extraño pensar que los debates morales se limitan a eso. Además, si una persona dice "*x* está bien" mientras que un miembro de otra sociedad dice "*x* no está bien", entonces el relativista cultural debe decir que no están en desacuerdo entre sí, ya que cada uno sólo está informando de las convenciones de su propia sociedad.[6] Incluso si digo: "No, no sólo quiero decir que los miembros de mi sociedad no deberían hacer *x*; quiero decir que la gente de tu sociedad tampoco debería hacer *x*."Las prácticas de tu sociedad están mal"."Las prácticas de tu sociedad son erróneas", no por ello estoy en desacuerdo con el extranjero que dice que *x* es correcto. Porque al decir que las prácticas de la sociedad extranjera son "moralmente incorrectas", según la teoría (c), todo lo que podría querer decir es que entran en conflicto con las normas de mi sociedad. Pero también en esto

se equivoca el relativista: cuando digo que las acciones nazis fueron malas, obviamente no estoy diciendo simplemente que los nazis tuvieran convenciones divergentes de las nuestras, ni estoy diciendo algo ininteligible.

En tercer lugar, aunque la teoría (c) no implica que ningún individuo sea moralmente infalible, sí implica que la sociedad es moralmente infalible. Según (c), todo lo que la sociedad respalda es correcto, automáticamente. Por lo tanto, todo el mundo debería aceptar observaciones como la siguiente:

> Si nuestra sociedad aprueba el aborto, entonces el aborto es correcto.
>
> Si nuestra sociedad no aprueba alimentar a los sin techo, entonces alimentar a los sin techo no es correcto.
>
> Si nuestra sociedad aprueba torturar a los bebés, entonces torturar a los bebés está bien.

No creo que ninguna de esas afirmaciones sea cierta, y tú tampoco deberías. Incluso si usted está a favor del aborto, no creo que deba respaldar la afirmación de que si la sociedad aprueba el aborto, entonces, automáticamente, es correcto.[7] Nuestra concepción moral ordinaria no considera la aprobación social como algo decretario.

En cuarto y último lugar, ¿por qué debemos obedecer las costumbres sociales? O bien hay buenas razones para las costumbres -es decir, razones que demuestran que las costumbres o el comportamiento que respaldan son buenos de alguna manera- o no hay tales razones. Si hay tales razones, entonces existen al menos algunos hechos evaluativos previos a las costumbres. Si no hay tales razones, entonces las costumbres son simplemente reglas arbitrarias, y ¿por qué deberíamos obedecer reglas arbitrarias?

Para ilustrar el dilema, supongamos que el relativista dice que las costumbres existen para promover la estabilidad y la supervivencia de los grupos sociales. Pues bien, ¿son buenas la estabilidad y la supervivencia de los grupos sociales? Si es así, entonces ese es al menos un hecho evaluativo que es independiente de las costumbres. Pero si no es así, entonces la existencia de una costumbre no crea ninguna razón para actuar de la manera que la costumbre requiere. Se puede decir que las costumbres crean razones no reales para las acciones -quizás debas obedecerlas para evitar el castigo de otras personas-. Pero, a menos que se plantee la obligación de evitar el castigo, esto no mostraría por qué una persona está moralmente obligada a seguir las costumbres sociales, sino sólo por qué es prudente hacerlo. Y es muy posible que uno pueda salirse con la suya violando las costumbres en algunas ocasiones, en cuyo caso el presente punto de vista no nos da ninguna razón para actuar "moralmente".

Ahora me referiré a dos respuestas que los relativistas pueden dar a algunos de estos problemas. La primera respuesta la llamo "el movimiento de rigidización". Tradicionalmente, los relativistas culturales han sido acusados de respaldar afirmaciones como:

> Si la sociedad aprobara comer niños, entonces comer niños sería bueno.

lo cual es claramente falso. El Movimiento Rigidizador afirma que la palabra "bueno" al final de la frase anterior se refiere a lo que la sociedad aprueba actualmente, de hecho, (a diferencia de lo que aprobaría en la situación hipotética descrita al principio de la frase).[8] Dado que la sociedad no aprueba actualmente el consumo de niños, la frase en su conjunto es falsa. Para ilustrar la idea, imagine que estoy en Boulder cuando pronuncio la siguiente frase: "Si hubiera volado a Minneapolis hoy, entonces Minneapolis habría estado aquí". Esta frase es falsa.

Minneapolis seguiría estando donde está; no se habría trasladado de alguna manera a Colorado. El término "aquí" que aparece al final de la frase citada se refiere al lugar en el que el hablante se encuentra de hecho, no al lugar en el que habría estado en la situación hipotética descrita. Según el movimiento rigidizador, la palabra "bueno" funciona como "aquí". Por lo tanto, el relativista no está comprometido con la absurda afirmación de que si aprobamos el canibalismo, entonces el canibalismo sería bueno.

Mi respuesta es que el relativista sigue comprometido con la aprobación:

Si la sociedad aprueba comer niños, entonces comer niños es bueno.

Esta ligera modificación -convertir la frase ofensiva del modo subjuntivo al indicativo- es suficiente para derrotar el movimiento rigidizador. Invocando de nuevo la analogía con la palabra "aquí", consideremos la frase "Si he volado a Minneapolis (de modo que ahora estoy en Minneapolis), entonces Minneapolis está aquí". Esta frase es verdadera, esté donde esté el hablante cuando la diga. Por tanto, la propuesta de que "bueno" funciona de forma análoga a "aquí" no ofrece esta vez ningún consuelo al relativista.[9]

La segunda respuesta a favor del relativismo la llamo la respuesta de los "Valores Fundamentales", y se aplica a los problemas del nazismo y la falibilidad. La respuesta de los valores fundamentales sostiene que algunas prácticas socialmente aceptadas pueden ser moralmente incorrectas en virtud de su falta de cohesión con otras prácticas sociales más fundamentales. Por ejemplo, supongamos que la sociedad antillana primitiva aceptaba una norma general de respeto a las personas pero que, debido a falsas creencias fácticas, no clasificaba a los negros como personas. En ese caso, el relativista puede argumentar que la esclavitud

era mala en la América primitiva, a pesar de su aceptación social: era mala en virtud de su conflicto con la norma más fundamental del respeto a las personas.[10]

Esta respuesta no contribuye a reducir el carácter desagradable del relativismo. El punto de vista sigue implicando que una sociedad es infalible en sus valores fundamentales. Además, no hay ninguna razón para suponer que la aceptación de la esclavitud en el pasado se debió a errores de hecho (no morales). Tal vez, en cambio, la aceptación de la esclavitud en el pasado se debió a una antipatía fundamental que se sentía hacia los miembros de otras razas, sin otra razón que la de pertenecer a otras razas. Tal vez los argumentos sobre las capacidades intelectuales inferiores de los africanos fueran racionalizaciones poco sinceras de esta hostilidad racista subyacente. Y quizás esta antipatía era tan profunda e intensa que habría anulado cualquier otro valor que entrara en conflicto con ella. No sé si este es el caso. Pero no parece que este sea el caso y que la institución de la esclavitud en la América primitiva sea irreprochable. En todo caso, esto la haría aún más reprobable. Sin embargo, la respuesta de los Valores Fundamentales implica lo contrario. Según el relativista, siempre que las actitudes racistas sean suficientemente fundamentales, escapan a toda posible crítica moral.

3.4 La teoría del mandato divino

A menudo oímos advertencias en el sentido de que el declive de las creencias religiosas socava la moralidad. Hay más de una forma de interpretar esta preocupación, pero aquí me centraré en la sugerencia de que no podría haber verdades morales si Dios no existiera: como dice Dostoievski, si no hay Dios, todo está permitido.[11]

La Teoría del Mandato Divino de la ética sostiene que una acción es moralmente correcta si y sólo si es del tipo que Dios ordena (o aprueba, o quiere que realicemos).[12] La versión de la teoría que quiero discutir sostiene una opinión similar sobre todas las demás propiedades evaluativas, incluyendo la bondad, la justicia, etc., es decir, que todas estas propiedades dependen de Dios de tal manera que nada podría tener ninguna propiedad evaluativa si Dios no existiera; sin embargo, en lo que sigue, me centraré en la propiedad de la rectitud. Dado que considera que la rectitud es reducible y depende de las actitudes de un observador (Dios), la Teoría del Mandato Divino es una forma de subjetivismo. Merece la pena señalar esto, ya que la teoría se considera a menudo la archienemiga del relativismo cultural, cuando en realidad ambas son variantes del mismo planteamiento metaético básico.

Nótese que la teoría no se limita a decir que siempre es correcto obedecer los mandatos de Dios, o que podemos averiguar lo que es correcto consultando los mandatos de Dios. Estos puntos de vista serían consistentes con la idea de que las acciones correctas son correctas independientemente de Dios, pero que Dios (siendo omnisciente) siempre sabe qué acciones son correctas y (siendo todo bueno) siempre las aprueba. En cambio, la teoría del mandato divino sostiene que las acciones correctas son correctas sólo porque Dios las ordena.

¿Tiene la teoría del mandato divino los mismos problemas que otras formas de subjetivismo? La respuesta es que hay dos problemas que son peculiares de la teoría del mandato divino; por otro lado, la teoría evita al menos un problema que afecta al subjetivismo secular. Consideremos seis problemas potenciales.

En primer lugar, ¿existe Dios? Si no existe Dios, la teoría del mandato divino es poco atractiva. Otras teorías subjetivistas evitan este problema, ya que no hay duda de que los individuos y las sociedades

humanas existen. Hay motivos serios para dudar de la existencia de Dios; sin embargo, están fuera del alcance de este libro.

En segundo lugar, ¿cómo sabemos lo que Dios aprueba? De nuevo, este problema no tiene análogos en las formas seculares de subjetivismo, ya que es fácil saber lo que aprueban los seres humanos. Según la tradición religiosa occidental, podemos conocer los deseos de Dios a partir de textos sagrados como la Biblia o el Corán. Pero, de hecho, esta fuente de orientación es notoriamente poco fiable. La guía moral que se encuentra en la Biblia incluye los siguientes imperativos (mis paráfrasis):

> Mata a cualquiera que maldiga a sus padres. (*Levítico 20:9*)
>
> Matar a cualquiera que cometa adulterio. (*Levítico 20:10*)
>
> Matar a los homosexuales. (*Levítico 20; 13*)
>
> Matar a las mujeres que tengan relaciones sexuales prenupciales. (*Deuteronomio 22:20-1*)
>
> Matar a las personas que trabajan en el día de reposo. (*Éxodo 35:2*)
>
> Hacer la guerra a los pueblos que ocupan la tierra prometida. No tengas piedad; mata a todo hombre, mujer y niño. (*Deuteronolny 7:1-2; 20:16-17*)
>
> La esclavitud está bien (*Levítico 25:44-5*)
>
> Está bien golpear a tus esclavos, siempre y cuando no mueran. (*Éxodo 21:20-1*)

Para no ser provinciano, observo que el Corán nos aconseja:

> Hacer la guerra y matar a los que no aceptan el Islam. (*Sura 9.5, 9.29-31*)
>
> Quien muera luchando por Alá irá al cielo. (*47.4-6*)
>
> Los incrédulos son las criaturas más viles. (*8.55*)
>
> Sufrirán horribles torturas en el infierno, para siempre. (*22.19-22*)

> Los castigos apropiados por oponerse al Islam incluyen la ejecución, la crucifixión, el corte de manos y pies y el exilio. (*5.33*)
>
> Cortar las manos a los ladrones. (*5.38*)

Pocos hoy en día, incluso entre los creyentes, defenderían las ideas y preceptos anteriores. Una alternativa obvia para algunos teístas sería sostener que, aunque existe un Dios, Él no inspiró la Biblia ni el Corán. El problema es que esto dejaría en el aire el misterio de cómo podemos saber lo que Dios quiere, amenazando así nuestra capacidad de conocimiento moral. Pero parece claro que tenemos algún conocimiento moral, como casi cualquier teísta estaría de acuerdo.

Otra alternativa, para un cristiano o un judío, sería sostener que la Biblia contiene algunos errores introducidos por los transmisores humanos de la palabra de Dios, que pueden haber malinterpretado partes del mensaje de Dios, insertado algunos de sus propios prejuicios o cometido errores al transmitir las escrituras a lo largo de las generaciones. Esto permitiría rechazar sistemáticamente los preceptos que he enumerado más arriba, manteniendo al mismo tiempo que la Biblia contiene muchas cosas verdaderas, importantes e inspiradas por Dios. No puedo refutar esta posición (no aquí, en todo caso), tampoco pretendo que no haya nada de valor en la Biblia o el Corán. Pero fíjese que si uno adopta esta posición, no puede simplemente apelar a la Biblia para determinar lo que está bien y lo que está mal. El teórico del Mandato Divino tendría que plantear algún tipo de acceso a los hechos morales independiente de nuestro conocimiento de la voluntad de Dios,[13] lo que parecería renunciar a una de las principales ventajas de la teoría del Mandato Divino, a saber, su capacidad para explicar el conocimiento moral. Por ejemplo, el teórico del Mandato Divino podría verse impulsado a plantear una facultad de intuición moral, pero entonces no está claro cómo su teoría sería mejor que el intuicionismo ético

tradicional. O el teórico podría adoptar una de las propuestas de los naturalistas éticos recientes (véase el capítulo siguiente, donde también se critican estas propuestas).

En tercer lugar, el problema de los nazis no aflige al teórico del mandato divino como aflige a otros subjetivistas: el teórico del mandato divino no necesita conceder la verdad de las afirmaciones morales nazis, ya que los nazis no actuaron de acuerdo con la voluntad de Dios. Sin embargo, se plantea un problema similar. El teórico del mandato divino debe sostener que si Dios nos mandara, sin ninguna razón en particular, torturar y asesinar a nuestros hijos, entonces tales acciones serían morales.

Un defensor de la teoría podría observar que, dado que Dios es moralmente perfecto, nunca ordenaría tales cosas.[14] Pero esto no responde a la objeción. Aunque sepamos que alguien no haría *A*, podemos preguntar qué pasaría si lo hiciera. Por ejemplo, no voy a beber una botella entera de Everclear. Sin embargo, podemos hablar de lo que sucedería si lo hiciera; por ejemplo, si me bebiera una botella de Everclear, debería llamar inmediatamente al hospital. Del mismo modo, podemos hablar de lo que sucedería si Dios nos ordenara asesinar niños, aunque estemos seguros de que no lo hará. Además, el razonamiento de la afirmación de que Dios nunca daría esa orden es falaz. Según la teoría del mandato divino, todo lo que Dios ordena es correcto, así que si Dios ordenara el asesinato, el asesinato sería correcto. Por lo tanto, la perfección moral de Dios no es un obstáculo para que emita tal orden: un ser moralmente perfecto no está impedido de ordenar a la gente que haga cosas que son moralmente correctas en el momento en que las ordena.[15]

Robert Adams propone otra respuesta al problema. Dice que la obligación de obedecer los mandatos de Dios está supeditada a que exista un Dios amoroso. Si Dios ordenara matar a los niños, esto

demostraría que no es amoroso como los cristianos creen que es de hecho. Dado que la rectitud de una acción consiste en su conformidad con la voluntad de un Dios amoroso, el asesinato de niños no sería correcto en esa situación.[16] Pero tampoco, según la opinión de Adams, estaría mal. Dado que, según Adams, el hecho de que un acto sea ordenado (prohibido) por un Dios amoroso es una condición necesaria para que sea correcto (incorrecto), Adams no tiene más que decir que si Dios no fuera amoroso, si se abstuviera de prohibir el asesinato de niños, o si simplemente no existiera, entonces no habría nada malo en matar niños. Esto es apenas más plausible que la posición de que el hecho de que Dios ordene el asesinato de niños lo haría correcto.

En cuarto lugar, la teoría del mandato divino está menos sujeta al problema del desacuerdo que otras fuentes de subjetivismo, porque no es absurdo suponer que las personas que no están de acuerdo con las cuestiones éticas a veces no están de acuerdo con lo que Dios aprobaría. Sin embargo, parece extraño sostener que todos los que discuten sobre cuestiones morales, incluyendo a los ateos, agnósticos y politeístas, están discutiendo sobre lo que Dios aprobaría.[17]

En quinto lugar, hay un problema que afecta a otras formas de subjetivismo y que no se aplica a la teoría del mandato divino. Se trata del problema de que los subjetivistas deben sostener que algún individuo o grupo es moralmente infalible, lo cual es inverosímil. Pero en el caso de Dios, no es inverosímil -de hecho, es cierto por definición- que sea infalible.

En sexto lugar, y lo más importante, incluso si hay un Dios, ¿por qué deberíamos hacer lo que Él dice? Que alguien te diga que hagas algo no crea en general una obligación moral de hacerlo. Si tu vecino te dice que mates a tu hijo,[18] esto no crea ninguna obligación, ni siquiera *prima facie*, para que lo hagas. Si Satanás te dice que no comas cerdo, esto no crea

ninguna obligación de obedecer. Debe haber algo especial en Dios, si sus mandatos crean obligaciones morales. ¿De qué se trata?

Yo sostengo que el teórico del mandato divino no tiene una respuesta satisfactoria a esta pregunta. En resumen:

1. Si ninguna característica de Dios fundamenta la obligación de obedecer los mandatos de Dios, entonces no hay obligación de obedecer los mandatos de Dios.
2. Las características moralmente neutras de Dios no fundamentan una obligación de obedecer los mandatos de Dios.
3. Si las características moralmente significativas de Dios fundamentan la obligación de obedecer sus mandatos, entonces algunos hechos morales son independientes de los mandatos y actitudes de Dios.
4. Si (a) no hay obligación de obedecer los mandatos de Dios o (b) algunos hechos morales son independientes de los mandatos y actitudes de Dios, entonces la teoría del mandato divino es falsa.
5. Por lo tanto, la teoría del mandato divino es falsa.

La premisa (1) ya ha sido motivada. Si no hay nada especial en Dios que lo distinga de los demás seres, que explique por qué debemos obedecerle en particular, entonces, puesto que en general no estamos obligados a obedecer a otros seres, no estamos obligados a obedecer a Dios.

En las premisas (2) y (3), una característica "moralmente neutra" es aquella que es posible tener sin tener una propiedad evaluativa; una característica moralmente significativa es aquella que no se puede tener sin tener una propiedad evaluativa. Por ejemplo, ser poderoso y ser el creador de la Tierra son características moralmente neutras. Ser justo y ser cruel son características moralmente significativas, ya que no se puede ser justo sin ser bueno en algún aspecto, ni se puede ser cruel sin

ser malo en algún aspecto.

(2) puede apoyarse en ejemplos hipotéticos. Para cualquier característica moralmente neutra de Dios, es posible imaginar que un ser malvado -Satán, por ejemplo- tenga esa característica. En tal situación, intuitivamente, no habría obligación moral de seguir las órdenes de Satanás. Por ejemplo, se podría decir que debemos obedecer a Dios porque es nuestro creador. Pero imagina que descubres que en realidad has sido creado especialmente por Satanás. En ese caso, ¿tendrías la obligación general de seguir la voluntad de Satanás? También se dice que Dios tiene un poder ilimitado sobre nosotros, y puede otorgar castigos y recompensas eternas. Imagina que Satanás tuviera un poder ilimitado sobre ti, y pudiera otorgar castigos y recompensas eternas. En ese caso, sin duda sería prudente obedecer las órdenes de Satanás. Pero, ¿sería moralmente obligatorio hacerlo; sería usted inmoral si se resistiera a los planes malvados de Satanás? Por último, supongamos que Satanás fuera omnisciente y omnipotente (pero aún así totalmente malvado). ¿Sería entonces moralmente obligatorio seguir su voluntad? En todos estos casos, la respuesta es seguramente que no estarías obligado a seguir la voluntad de Satanás. Por lo tanto, ninguna de estas propiedades -ser nuestro creador, tener la capacidad de recompensarnos y castigarnos, ser omnisciente, ser omnipotente, o cualquier combinación de ellas- puede fundamentar una obligación de obediencia por nuestra parte hacia otro ser.

Hay un aspecto crucial, o familia de aspectos, en el que Dios difiere de Satanás en mis escenarios hipotéticos. Se trata de que Satanás es malo, mientras que Dios es puramente bueno, justo, benévolo, etc. Es plausible que si un ser supremamente bueno te dice que hagas algo, entonces estás obligado a hacerlo. Pero aquí es donde entra la premisa (3). Si la razón por la que debemos obedecer a Dios se basa en la bondad de Dios,

entonces deben existir algunos hechos morales previos a las órdenes de Dios. Por hipótesis, la capacidad de los mandatos de Dios para generar verdades morales depende de la bondad de Dios; por tanto, so pena de circularidad, no puede ser también el caso que la bondad de Dios sea generada por sus mandatos. (No puede ser, por ejemplo, que Dios sea bueno porque se aprueba a sí mismo, o porque nos ordena adorarle. Imaginemos que Satanás también se aprueba a sí mismo y nos ordena que le adoremos). Dios ya debe ser bueno, por lo que ya debe haber una norma de valor -por ejemplo, que el amor, la misericordia y la justicia son buenos- antes de los mandatos de Dios.

A propósito de esto, recordemos la solución de Robert Adams al problema de la posible orden de Dios de matar niños. Adams considera que el amor de Dios es relevante para la obligación de obedecerle y argumenta que un Dios amoroso no podría ordenar esa matanza. Pero, ¿qué motiva a Adams a destacar el amor? ¿No será que considera este rasgo como una virtud, o de alguna manera como algo bueno?[19] Es difícil imaginar a alguien que no valorara el amor dando la respuesta de Adams. Así pues, parece que la capacidad de los mandatos de un Dios amoroso para crear obligaciones depende del valor asumido independientemente del amor.

Por último, la premisa (4) es verdadera porque la teoría del mandato divino sostiene que lo que es moralmente obligatorio lo es en virtud de que Dios lo ordena, y que algo similar ocurre con todas las demás propiedades morales (como la bondad y la maldad). Esto implica tanto que es obligatorio obedecer los mandatos de Dios, como que nada tiene ninguna propiedad moral independientemente de los mandatos de Dios.

La conclusión (5) se deriva de las premisas (1)-(4): Las características no morales de Dios no pueden fundamentar una obligación de obediencia. Las características morales de Dios pueden

fundamentar tal obligación sólo si hay hechos morales independientes de los mandatos de Dios. Por lo tanto, o bien no hay ninguna razón para obedecer a Dios, o bien existen hechos morales independientes de los mandatos de Dios. Cualquiera de estas alternativas entra en conflicto con la teoría del mandato divino. Así que esa teoría debe ser rechazada.

Un punto estrechamente relacionado puede hacerse más sucintamente. ¿Por qué Dios ordena lo que hace? Si Dios no tiene razones morales para sus mandatos, entonces son meramente arbitrarios, y ¿por qué deberíamos obedecer mandatos arbitrarios? Pero si Dios tiene razones morales para sus mandatos, entonces algunas verdades morales deben existir independientemente de sus mandatos. En cualquier caso, la teoría del mandato divino es falsa.[20]

Una respuesta a esto es que los mandatos de Dios pueden escapar a la arbitrariedad en virtud de que Dios tiene razones no morales para sus mandatos.[21] No creo que esto ayude, al igual que la sugerencia de que Dios tiene malas razones para sus mandatos. ¿Cómo podría servir el hecho de que Dios tenga una razón sin peso moral para un mandato para hacer que ese mandato no sea arbitrario en un sentido moralmente relevante, es decir, de tal manera que debamos seguirlo moralmente?

*3.5 La teoría del observador ideal

Las teorías del observador ideal dicen que una cosa es buena, o una acción correcta, si un "observador ideal" adoptaría una actitud favorable a ella.[22] Esto es similar a la teoría del mandato divino (Dios es el último observador ideal), pero no requiere que exista realmente ningún ser divino; en su lugar, debemos responder a las cuestiones morales razonando hipotéticamente sobre lo que favorecería el observador ideal, si lo hubiera. Clasifico esto como una forma de subjetivismo, ya que

hace que el valor dependa constitutivamente de las actitudes de un observador (hipotético).[23]

Este estado de la teoría plantea tres cuestiones. En primer lugar, ¿qué tipo de "actitud favorable" se supone que adopta el observador ideal? En segundo lugar, ¿qué se entiende por "observador ideal"? Y tercero, ¿es el bien lo que todos los observadores ideales favorecerían necesariamente, lo que un observador ideal podría favorecer, o qué?

No me detendré en la primera pregunta. La actitud favorable podría ser el deseo, la aprobación o cualquier otra cosa. En lo que sigue, hablaré en términos de deseo, pero se pueden formular argumentos análogos independientemente de la actitud favorable que utilice el teórico del observador ideal.

En respuesta a la segunda pregunta, he aquí algunos rasgos que se han atribuido al "observador ideal":

> Conoce todos los hechos no morales.
>
> Imagina todos los estados de cosas relevantes de forma perfectamente vívida.
>
> Es perfectamente racional.
>
> Es desinteresado y desapasionado, es decir, carece de deseos y emociones que se refieran a individuos concretos.
>
> Es benévolo.
>
> Es normal en otros aspectos (por ejemplo, no está loco).[24]

Diferentes filósofos pueden elegir conjuntos de rasgos ligeramente diferentes, lo que conduce a diferentes versiones de la teoría del observador ideal. No hay que preocuparse por esto. Sin embargo, hay dos puntos importantes sobre lo que no se puede considerar que es el observador ideal: Al observador ideal no se le debe asignar ninguna

propiedad moral; por ejemplo, no debe entenderse como un ser perfectamente virtuoso. Esto llevaría a la teoría a la circularidad, dado su objetivo de explicar todas las propiedades morales. Además, aunque el observador ideal pueda tener plena conciencia de los hechos no morales, no se le debe atribuir conocimiento moral, ni ningún rasgo que implique una capacidad de conocimiento moral (como una facultad de intuición ética). De nuevo, la preocupación sería de circularidad; tal capacidad presupondría la existencia de hechos morales independientes del observador.

En respuesta a nuestra tercera pregunta, hay dos tipos de teoría del observador ideal. Una toma el bien como aquello que todo observador ideal aprobaría necesariamente.[25] El problema es que esto parece dar lugar a que nada sea bueno. Los rasgos enumerados anteriormente (aparte de la benevolencia) no implican que el observador así descrito tenga ningún deseo, y mucho menos sentimientos de aprobación o desaprobación. Podríamos apelar a la condición de "normalidad" para tratar esto: un ser sin deseos sería anormal. Tal vez haya algunos deseos que todas las personas normales deben tener. Ahora bien, el problema es que la definición del observador ideal parece incluir ilícitamente un término evaluativo: cuando "normal" se utiliza para decir "tiene deseos normales", es evaluativo. Por ejemplo, las personas que piensan que la homosexualidad está mal se niegan a llamar "normales" a los deseos homosexuales; los que la consideran permisible llaman a esos deseos "normales". Llamamos 'anormal' a la pedofilia porque es mala. El defensor de la teoría del observador ideal probablemente querría excluir a los observadores con deseos sádicos; si esto se hace sobre la base de la "anormalidad" de tales deseos, esto parecería ser porque "anormal" funciona evaluativamente. Un punto similar se aplicaría al intento de utilizar el requisito de "racionalidad" para descartar a los observadores ideales apáticos.

Podríamos intentar simplemente estipular qué deseos tiene el observador ideal; por ejemplo, que sea benevolente (que desee el bienestar de la gente en general).[26] Un problema con esto es que, al estipular las motivaciones del observador ideal, estaríamos simplemente estipulando el contenido de la teoría moral. Si decimos: "Los observadores ideales desean *x*, *y* y *z*", entonces estamos estipulando que *x*, *y* y *z* son los únicos bienes. Esto se debe a que nada más en la definición del observador ideal implica ninguna motivación, por lo que las únicas cosas que todos los observadores ideales desearían serían *x*, *y* y *z*. Pero entonces podríamos también tomar la teoría más simple, "Para que una cosa sea buena es que sea *x*, *y* o *z*", y llamarnos naturalistas éticos. No se hace ningún trabajo añadiendo el resto del aparato del observador ideal.

Richard Brandt ofrece un enfoque mejor para nuestra tercera pregunta: relativiza el valor al individuo, y dice que lo que es bueno para una persona determinada es lo que esa persona desearía, si fuera plenamente racional, estuviera plenamente informada, etc.[27] Esto nos permite partir de nuestros deseos reales, y simplemente considerar qué cambios se harían si llegáramos a ser plenamente racionales. Es de suponer que para llegar a ser plenamente racional no tendría que perder todos mis deseos, por lo que algunas cosas serán buenas (en relación a mí) según esta teoría.

Todo eso fue a modo de identificar lo que considero la mejor versión de la teoría del observador ideal. Ahora, lo que está mal en ella.

Los problemas son conocidos. Primero, el problema de los deseos horribles: Supongamos que quiero ver mucha muerte y destrucción porque sí. Imaginar vívidamente la destrucción sólo aumenta mi deseo de ella. El deseo no ha sido causado por ninguna creencia falsa mía, ni me han lavado el cerebro para ello, etc.; simplemente soy una persona

desagradable.[28] Ninguna otra información no moral me haría cambiar de opinión al respecto. Según la Teoría del Observador Ideal, parece que la muerte y la destrucción son buenas (para mí), y debo apuntar a ellas. Brandt podría responder que ninguna persona persona es así. No veo ninguna razón para suponerlo, pero no es necesario debatirlo. He aquí un ejemplo más realista: hay un individuo que tiene una aversión básica a las personas de razas de piel oscura. Si se le diera toda la información no moral, pero ningún argumento o intuición moral, seguiría sintiendo esta aversión. La aversión, supongamos, es una hostilidad programada genéticamente que algunas personas sienten por las personas de otras razas y grupos sociales. En esta situación, la teoría del observador ideal aconsejaría aparentemente que esta persona favoreciera las acciones y políticas racistas, quizás incluso la esclavitud para los miembros de la raza despreciada.

Brandt podría afirmar que los racistas reales no son así, que las actitudes racistas se basan de hecho en falsas creencias no morales sobre otras razas. Veo pocas razones para creerlo; al menos es igualmente plausible que las creencias racistas sobre los rasgos de personalidad y las capacidades intelectuales y de otro tipo de las razas minoritarias se acepten como racionalizaciones de sentimientos racistas preexistentes. En cualquier caso, se admite que la teoría del observador ideal sostiene que si las actitudes racistas son fundamentales y no están causadas por errores de hecho, lavado de cerebro, etc., entonces las políticas racistas son buenas (para los racistas), y las personas con esas actitudes deberían buscar el trato desigual de las minorías.

En segundo lugar, el problema del desacuerdo: Según la versión de Brandt de la teoría del observador ideal, las personas que parecen estar en desacuerdo sobre la moral no están realmente en desacuerdo. Esto es cierto independientemente de que la mayoría de las personas sean lo

suficientemente parecidas como para que las versiones idealmente racionales de ellas deseen las mismas cosas. Es cierto en virtud de la forma lógica de los enunciados morales según la teoría: cada persona está hablando de lo que desearía si estuviera plenamente informada, etc. Por tanto, "*x* es bueno" y "*x* no es bueno" no se contradicen, si lo dicen personas diferentes.

Tercero, el problema de la falibilidad: la teoría del observador ideal puede explicar cómo mis preferencias reales pueden ser malas, ya que pueden no coincidir con las que tendría si estuviera plenamente informado", etc. Lo que no puede explicar es el hecho de que incluso una persona plenamente informada puede ser moralmente falible: la corrección en todas las cuestiones no morales no excluye conceptualmente el error moral. Si existiera un observador ideal, podría preguntarse inteligentemente: "¿Son realmente buenas las cosas que quiero?".

En cuarto y último lugar, el problema de la arbitrariedad: ¿Por qué el observador ideal desea lo que hace? Si lo desea porque es bueno, o por cualquier otra razón moral, entonces hay hechos morales independientes de sus actitudes. Pero si no tiene ninguna razón moral para desear lo que hace, entonces sus deseos son arbitrarios y no tienen ninguna importancia moral. ¿Por qué el hecho de que alguien desee o apruebe algo sin razón alguna hace que esa cosa sea buena? Esto es esencialmente una apelación a la intuición ética de que el hecho de ser deseado no convierte un objeto moralmente neutro o malo en bueno (no confundir con la afirmación de que la satisfacción del deseo no es buena).

3.6 La falacia subjetivista

¿Por qué alguien sostiene una teoría subjetiva del valor? Los subjetivistas individuales y los relativistas culturales suelen intentar respaldar sus puntos de vista con uno o ambos de los siguientes tipos de pruebas: (*i*) pruebas de que las creencias morales varían entre individuos o culturas; (*ii*) pruebas de que las creencias morales están influidas por las emociones, la presión de los compañeros y otras causas no racionales. El argumento parece ser:

1. Las diferentes personas/culturas tienen muchos conjuntos diferentes de creencias morales.
2. Por lo tanto, hay muchas verdades morales diferentes (muchas perspectivas morales diferentes son igualmente correctas).

O:

3. Las creencias morales de las personas son generalmente causadas por las emociones y/o las convenciones de su sociedad.
4. Por lo tanto, una creencia moral es sólo una creencia sobre las emociones y/o las convenciones.

Es difícil creer que alguien necesite que se le señalen las falacias de lo anterior. El primer argumento es inválido tal y como está; para que (1) apoye a (2), habría que añadir la premisa "lo que la gente cree es cierto". Esa premisa es falsa, y se sabe que es falsa. El segundo argumento es igualmente inválido; para que (3) apoye a (4), habría que añadir la premisa: "Las creencias de la gente son siempre causadas por lo que son". También se sabe que eso es falso. La gente se equivoca a menudo y suele adoptar creencias por razones irracionales.

Los argumentos de este tipo merecen la etiqueta de "falacia subjetivista", ya que son prácticamente los únicos argumentos que dan los subjetivistas, y porque los dan casi todos los subjetivistas.[29] El error

subyacente es sencillo: consiste en confundir las representaciones con la realidad. El primer argumento anterior confunde abiertamente las creencias con las verdades. El segundo confunde la forma en que la gente forma una creencia con lo que la creencia es.

Para ilustrarlo, supongamos que tres jugadores en el hipódromo piensan, por razones puramente emocionales, que un caballo diferente va a ganar. No se deduce que la carrera vaya a tener tres resultados diferentes, uno en el que cada apostante tenga razón. Tampoco se deduce que decir que un caballo va a ganar la carrera sea sólo decir que se tiene una determinada emoción. Sin embargo, ese es el tipo de razonamiento que implica la Falacia Subjetivista.

En resumen, las formas más comunes de subjetivismo están motivadas por confusiones obvias. Además, todos los subjetivistas éticos se enfrentan al menos a tres problemas importantes:

i) *La probabilidad de actitudes horribles*: Los subjetivistas deben sostener que acciones aparentemente horribles, como el asesinato y la tortura, son moralmente correctas siempre que la persona o el grupo apropiado las respalde.

ii) *El problema del desacuerdo*: Los subjetivistas no pueden explicar de forma plausible el desacuerdo moral.

iii) *El problema de la arbitrariedad*: Los subjetivistas deben sostener que las cosas se hacen buenas, o las acciones se hacen obligatorias, mediante reglas o actitudes meramente arbitrarias.

Además, los subjetivistas individuales, los relativistas culturales y los teóricos del observador ideal se enfrentan:

iv) *El problema de la falibilidad*: No pueden explicar la intuición de que la persona o el grupo en cuestión puedan tener actitudes moralmente incorrectas.

Y los teóricos del mandato divino se enfrentan a:

v) *El problema de la existencia de Dios*: Es dudoso que Dios exista.

vi) *El problema de conocer los deseos de Dios*: Aunque exista, no está claro qué quiere Dios.

Estas razones parecen más que suficientes para rechazar el subjetivismo. Pero, aunque me equivoque en todos estos problemas, los subjetivistas seguirán enfrentándose a los problemas que se plantean a todos los reduccionistas en el capítulo siguiente.

Notas

1. En realidad, el signo "=" se refiere a lo que los filósofos quieren decir cuando hablan de que algo es "reducible a" algo. No puede representar la sinonimia o la relación de definibilidad, ya que eso dejaría fuera a los reduccionistas sintéticos. No puede representar la identidad, ya que la identidad es simétrica, mientras que la reducibilidad se supone que es asimétrica: las cosas de la izquierda se reducen a las cosas de la derecha, pero las cosas de la derecha no se reducen a las cosas de la izquierda. La locución "... explica lo que es..." que aparece en el texto es mi glosa de esta desconcertante relación.

2. Sobre (b), véase Westermark (1960, pp. 114-16); Hume ([1739]1992, p. 469); y Hobbes ([165 IJ 1996, p. 39), aunque Hobbes sustituye el deseo por la aprobación. Sobre (c), véase Benedict (1934, p. 73); Harman (1977, p. 94); y Sumner (1907, pp. 29, 58-9, 418). Sobre (d), véase Quinn (2002) y Adams (1981; 1999).

3. Los problemas para el subjetivismo y el relativismo que involucran a los nazis, el desacuerdo y la falibilidad son bien conocidos y se presentan comúnmente en los cursos y libros de texto universitarios de filosofía. Véase James Rachels (2003, pp. 21-3, 35-6) para un ejemplo particularmente conocido y claro. El problema que implica el desacuerdo se atribuye a veces a G. E. Moore (1947, pp. 62-4), aunque se remonta al menos a Thomas Reid ([1788] 1986, pp. 367-8). que lo despliega con gran efecto contra Hume.

4. Con esto quiero decir que mis actitudes morales siguen la verdad moral de forma infalible, no que mis creencias sobre mis actitudes morales sean siempre correctas. Es

posible que me equivoque sobre lo que apruebo, aunque creo que incluso este tipo de error sería difícil de encontrar.

5. Goldhagen (1996) sostiene que el tisemitismo, hasta el punto de aprobar el asesinato de judíos, era un elemento central de la cultura alemana de la época, lo que llevó a cientos de miles de alemanes de a pie a participar en el Holocausto de buen grado. La tesis de Goldhagen es controvertida entre los historiadores (véanse los artículos de Shandley 1998), pero podemos suponer, en aras del argumento, que su versión de la historia es correcta. El punto filosófico relevante es que la verdad de la opinión de Goldhagen no haría menos malvada la participación alemana en el holocausto; si acaso, todo lo contrario.

6. G. E. Moore (1947, pp. 68-9).

7. Asumo que el "si-entonces" aquí es algo más fuerte que el condicional material. Incluso si el "si-entonces" en inglés a veces o siempre significa el condicional material, la teoría (c) claramente licencia la afirmación de algo más fuerte que un condicional material aquí, algo en la línea de que la aprobación del aborto por parte de la sociedad sería suficiente para hacerlo correcto.

8. Puntos análogos se aplican a los defensores de la teoría (b). El nombre "movimiento rigidizador" indica que el subjetivista propone tratar los términos evaluativos como designadores rígidos (véase Kripke 1972, p. 48). Vallentyne (1996, p. 105-9) discute la jugada rigidificadora como una forma de evitar problemas como la objeción de los nazis; sin embargo, Vallentyne toma esta jugada para convertir la teoría en una forma de objetivismo.

9. Esta es, por supuesto, la razón por la que formulé mi tercera objeción al relativismo utilizando condicionales indicativos en primer lugar.

10. Esta respuesta se deriva de Harman (1977, pp. 94-5).

11. Se trata de una paráfrasis de las opiniones de Iván Karamozov en Los hermanos Karamazov.

12. Véase Adams (1981); Quinn (2002); Alston (1990); pero nótese que la teoría del bien de Adams (1999, capítulo 1), que considera que la bondad consiste en la semejanza con Dios, no es estrictamente una forma de la teoría del mandato divino tal y como yo la defino, ya que Adams no considera que la bondad dependa constitutivamente de los mandatos o actitudes de Dios (o de nadie). Véase Carson (2000, pp. 239-56) para una variante que implica preferencias divinas, en lugar de órdenes.

13. Como propone Quinn (2002, p. 673). Una alternativa sería intentar, mediante la investigación histórica, determinar cuándo aparecieron por primera vez los errores en las versiones aceptadas de la Biblia. Pero dudo que el planteamiento de atribuir la errancia bíblica a un error humano dé los resultados deseados, como permitirnos condenar la esclavitud, a menos que supongamos que algunos de los receptores originales de la supuesta revelación cometieron ellos mismos errores sobre la voluntad de Dios.

l4. Quinn (2002, p. 676).

15. Estos puntos aparecen en Layman (1991, pp. 38-9).

16. Adams (1981, pp. 86-8). Esto es una ligera simplificación; Adams dice que el amor de Dios no se afirma explícitamente como parte del significado de un estado moral, sino que se presupone en los estados morales. Pero esta distinción no afecta a mi respuesta.

17. Aquí estoy pensando en la teoría del Mandato Divino como una forma de reduccionismo analítico, como creo que ha hecho la mayoría de los que ven una conexión necesaria entre la ética y Dios. Véase el capítulo 4 para más información sobre la distinción entre reduccionismo analítico y sintético.

18. Como supuestamente le dijo Dios a Abraham (Génesis 22:2). El contexto de la historia deja claro que Abraham debe ser alabado por su voluntad de seguir este mandato. Quinn (2002, p. 678) coincide en lo acertado de sacrificar a Isaac en tal situación.

19. Adams (1999, p. 281) lo confirma, afirmando que "la bondad de Dios" es "importante para acreditar los mandatos de Dios por su función de constituir una obligación".

20. Esto se llama a veces "el problema de Eutifrón" por su similitud con el problema planteado en el Eutifrón (Platón 1996, pp. 14-15).

21. Carson (2000, pp. 245-6) da esta respuesta, pero redactada en términos de preferencias y no de mandatos.

22. Véase Firth (1952) y Brandt (1979; 1996). Brandt (1979, pp. 225-8) distingue su punto de vista de la "teoría del observador ideal"; esto refleja una diferencia meramente terminológica entre Brandt y yo.

23. Firth (1952, pp. 322-4) no está de acuerdo con esta clasificación, ya que su teoría hace que el valor sea independiente de las actitudes de cualquier ser real; pero esto es una disputa meramente terminológica.

24. Esta lista es un compuesto de Firth (1952, pp. 333-45) y Brandt (1979, p. 14-15).

25. Firth (1952, p. 320).

26. Brandt (1979, pp. 227-11) discute esta propuesta pero no la respalda.

27. Brandt (1996, pp. 12,60, 113).

28. Brandt (1996, pp. 51-8) enumera varias formas en que un deseo puede ser irracional. No veo ninguna razón por la que mi deseo sádico imaginado deba ser irracional en cualquiera de esas formas. En cualquier caso, argumentaré en el capítulo 7 que Brandt se equivoca sobre lo que hace que los deseos sean irracionales.

29. Excepto los filósofos profesionales, como Harman, que generalmente son demasiado sofisticados para cometer la Falacia Subjetivista abiertamente. Pero es el único tipo de argumento que ofrecen Benedict (1934), Sumner (1907) y legiones de estudiantes universitarios.

4

Reduccionismo

En el presente capítulo, defiendo dos tesis principales, frente a las opiniones de los reduccionistas éticos: que al menos algunas propiedades evaluativas son irreductibles, y que el conocimiento moral no puede explicarse sin apelar a la intuición ética.

4.1 ¿Qué es el reduccionismo?

Tanto el naturalismo ético como el subjetivismo son formas de reduccionismo sobre la ética. El reduccionismo sostiene que todas las propiedades evaluativas son reducibles a propiedades naturales.

¿Qué significa esto? Considero que las "propiedades naturales" incluyen todas las propiedades (incluidas las propiedades relacionales) a las que se puede hacer referencia utilizando términos no evaluativos.[1] Esto incluye tanto las propiedades físicas (como el cristal, la forma y la carga eléctrica) como las mentales (como creer, sentirse triste y tener una sensación de cosquilleo). Los naturalistas y los subjetivistas discrepan entre sí sobre si ciertos tipos de propiedades mentales -en concreto, las actitudes de los observadores hacia los objetos de

evaluación- deben incluirse entre aquellas a las que se reducen las propiedades de valor; pero están de acuerdo en que las propiedades de valor se reducen a algún tipo de propiedades naturales. Decir que una propiedad evaluativa se "reduce" a algún conjunto de propiedades naturales es decir que se podría explicar la naturaleza de la propiedad evaluativa -explicar lo que es- únicamente en términos de esas propiedades naturales.

Inicialmente, el reduccionismo podría parecer paradójico: ¿cómo podrían utilizarse terminos totalmente no evaluativo para explicar la naturaleza del valor? Para ver que no es tan paradójico, consideremos una analogía: hay quien dice que para que un objeto sea rojo basta con que refleje predominantemente luz con longitudes de onda en torno a los 650 nanómetros. Obsérvese que esta explicación se da utilizando por completo términos no relacionados con el color ("reflejar", "longitud de onda", etc.), de modo que una persona sin conocimiento previo del color podría, si la explicación es correcta, llegar a comprender la naturaleza del color rojo a partir de ella. Sea o no correcta esta explicación de la naturaleza del color rojo, al menos no es obviamente falsa. Del mismo modo, tal vez sea posible dar una explicación de lo que es la bondad que no utilice ningún término evaluativo.[2]

Una de las principales ventajas que se atribuyen al reduccionismo es que promete hacer inteligible nuestro conocimiento de la moral. Si las propiedades morales son irreductibles, propiedades no naturales, entonces parece que la única forma de conocerlas es por "intuición", una facultad que los filósofos contemporáneos se han mostrado reacios a aceptar. Pero la opinión de que las propiedades morales son reducibles a las propiedades naturales mantiene la esperanza de que podamos conocer los hechos morales de una manera similar a como conocemos (otros) hechos naturales, como por ejemplo mediante la observación o el

razonamiento científico.[3]

En el resto de este capítulo, sostengo tanto que el reduccionismo sobre la ética es falso, como que esta supuesta ventaja de la teoría es ilusoria: los reduccionistas no pueden escapar a la apelación a la intuición como fuente de conocimiento moral.[4]

4.2 Reduccionismo analítico

El reduccionismo analítico sostiene que el significado de cualquier término moral puede darse utilizando términos no morales. Por ejemplo, un reduccionista analítico podría sostener que "*x* es bueno" es sinónimo de "*x* aumenta la cantidad total de disfrute en el mundo". O podría sostener que "*x* es bueno" es sinónimo de "*x* es algo que deseamos desear".[5]

El argumento de la pregunta abierta de G. E. Moore refuta el reduccionismo analítico.[6] Supongamos que "*N*" es una expresión que contiene sólo términos no valorativos que denotan alguna propiedad. ¿Son buenas las cosas que tienen *N*?" es una pregunta abierta. Lo que esto significa es que una persona podría dar coherentemente cualquier respuesta a la pregunta. Pero "¿Son buenas las cosas?" no es una pregunta abierta; una persona no podría dar coherentemente ninguna de las dos respuestas a esa pregunta. Por lo tanto, las dos preguntas no son equivalentes. Por lo tanto, "cosas buenas" no es sinónimo de "cosas que tienen *N*". Para ilustrarlo, consideremos la teoría de que "*x* es bueno" significa que "*x* aumenta la cantidad de disfrute en el mundo". Moore argumentaría:

1. "¿Es bueno hacer lo que aumenta la cantidad de disfrute en el mundo?" es una pregunta abierta.

2. "¿Es bueno hacer lo que es bueno?" no es una pregunta abierta.

3. Por lo tanto, las preguntas de (1) y (2) no son sinónimos.

4. Por lo tanto, "aumenta la cantidad de disfrute en el mundo" no es sinónimo de "es bueno".

La idea detrás de la premisa (1) es que una persona podría preguntarse coherentemente si aumentar el disfrute es bueno (después de todo, un sistema de valores ascético es lógicamente consistente). Pero como observa (2), uno no podría preguntarse si hacer el bien es bueno; ningún sistema de valores coherente niega que hacer el bien sea bueno.

La idea que subyace en (3) es que, puesto que una pregunta es abierta y la otra no, no son la misma pregunta. También se puede ver inmediatamente que las dos preguntas son diferentes, incluso sin invocar la noción de "preguntas abiertas".

El paso (4) se deduce de (3) porque las preguntas de (1) y (2) se diferencian únicamente en que una de ellas sustituye el predicado "es bueno" por el predicado "aumenta la cantidad de disfrute en el mundo"; por tanto, la diferencia de significado entre las dos preguntas debe atribuirse a una diferencia de significado entre esos dos predicados. Si los dos predicados tuvieran el mismo significado, como afirma el reduccionista, uno debería poder sustituir uno por otro en cualquier oración sin cambiar el significado de la misma.

Moore dice que un argumento similar funcionará contra cualquier teoría analítica reduccionista. Conozco una aparente excepción a esto: considere una teoría analítica del observador ideal, que sostiene que "*x* es bueno" significa algo así como "Un ser idealmente racional, conocedor y sensible aprobaría *x*". El argumento de la pregunta abierta puede fallar contra esta teoría, porque la primera premisa no es claramente verdadera.[7] Es decir, no está claro que preguntas como "¿Son

realmente buenas las cosas que aprobaría un ser idealmente racional, informado y sensible?" sean preguntas abiertas. Uno puede pensar que, siempre que caractericemos correctamente a los observadores ideales, un observador ideal está obligado a hacer las cosas bien.

Creo que hay dos maneras de entender este tipo de teoría, una que la hace no reduccionista e inmune al argumento de la pregunta abierta, y otra que la hace reduccionista pero vulnerable al argumento de la pregunta abierta. Por un lado, si se considera que la "racionalidad ideal" incluye la capacidad de hacer evaluaciones racionales, que el "conocimiento ideal" incluye el conocimiento moral, o que la "sensibilidad ideal" incluye la sensibilidad a las características morales de las cosas, entonces creo que "¿Las cosas que un ser idealmente racional, informado y sensible aprobaría son realmente buenas?" no es una pregunta abierta. Pero en ese caso, creo que tenemos una teoría no reduccionista, ya que la existencia de otros hechos o rasgos evaluativos debe presuponerse al definir "bueno". No me interesa en esta sección refutar los relatos no reduccionistas.

Si el conocimiento ideal sólo incluye el conocimiento no moral, y la sensibilidad ideal sólo incluye la sensibilidad no moral, entonces creo que es una pregunta abierta si un ser idealmente racional, conocedor y sensible también poseería creencias morales correctas; por lo tanto, es una pregunta abierta si lo que tal ser aprobara sería bueno. Es posible que un ser sea perfectamente racional, informado y sensible en todas las áreas no evaluativas y que, sin embargo, carezca de juicios morales correctos. En tal caso, este ser podría no aprobar lo que es bueno, o incluso aprobar cosas que son malas.

Creo, pues, que el argumento de Moore tiene éxito contra todas las teorías genuinamente reduccionistas de los significados de las expresiones evaluativas. Los reduccionistas suelen admitir que si el

argumento de la cuestión abierta tiene éxito contra teorías como la de que "bueno" significa "que conduce al placer", entonces tiene éxito contra todas las teorías analíticas reduccionistas. Sin embargo, han dicho que el argumento de Moore falla incluso contra la teoría de que "bueno" significa "conducente al placer", por al menos dos razones.

En primer lugar, hay quien dice que Moore plantea la pregunta. Dado que el reduccionista al que nos enfrentamos sostiene que el bien es simplemente lo que aumenta el placer, obviamente dirá que no es una cuestión abierta si es bueno hacer lo que aumenta el placer. Para Moore, afirmar lo contrario sólo supone que la teoría reduccionista es falsa.[8]

Esta respuesta emplea un concepto demasiado amplio de "plantear la pregunta". No es cierto que cada vez que un argumento despliega una premisa que contradice directa y obviamente la posición de un oponente, el argumento se hace la pregunta. Tampoco es cierto que siempre que un oponente coherente rechace al menos una de las premisas de un argumento, éste sea una incógnita. (Esta última condición se aplica a todo argumento válido). Consideremos otro famoso argumento filosófico: El filósofo *A* afirma que "conocimiento" significa "creencia justificada y verdadera". El filósofo *G* señala el siguiente tipo de ejemplo:

> Supongamos que Smith cree justificadamente que Jones es dueño de un Ford (ha visto a menudo a Jones conduciendo un Ford, ha visto el título del coche, etc.). Smith deduce correctamente que la siguiente afirmación también es cierta: "Jones tiene un Ford, o Brown está en Barcelona" (Barcelona ha sido elegida al azar; Smith no tiene ni idea de dónde está Brown). Pero por pura coincidencia, Brown resulta estar en Barcelona. En este caso, Smith cree que [Jones es dueño de un Ford, o que Brown está en Barcelona]; está justificado que lo crea (ya que lo infirió correctamente de una creencia justificada); y es cierto (ya que Brown está en Barcelona). Pero intuitivamente, Smith no sabe que [Jones es dueño de un Ford o que Brown está en

> Barcelona]. Por lo tanto, la creencia justificada y verdadera no es lo mismo que el conocimiento.

El argumento anterior se considera ampliamente, y con razón, que refuta de forma concluyente la explicación del conocimiento basada en la "creencia verdadera y justificada"; de hecho, es uno de los pocos ejemplos, y uno de los más célebres, de refutación concluyente de una opinión ampliamente defendida hasta ahora en la filosofía moderna.[9] Qué débil sería la respuesta de *A*:

> *G* no ha hecho más que plantear la pregunta. Yo digo que el conocimiento es una creencia justificada y verdadera. De esto se deduce directa y obviamente que Smith sabe que [Jones es dueño de un Ford o que Brown está en Barcelona] en el ejemplo de *G*. Que *G* afirme que Smith carece de tal conocimiento sólo supone que mi definición es errónea. Así que *G* no ha demostrado nada.

A está en lo cierto al observar que el preludio de *G* contradice directa y obviamente la teoría de *A*. Pero esto no significa que *G* se haya inventado la cuestión; sólo significa que la refutación de *A* por parte de *G* es directa y obvia. *G* tiene éxito mientras que *A* fracasa porque la tesis de *G*, una vez expuesta, es intuitivamente obvia, mientras que la teoría de *A* no es intuitivamente obvia, sino que necesita ser probada mediante ejemplos. Por tanto, *A* no está en condiciones de apelar simplemente a su teoría como justificación para negar la premisa que se utilizaría para refutar la teoría.

Del mismo modo, Moore tiene éxito frente al reduccionista porque es capaz de aducir premisas que, una vez enunciadas, son intuitivamente obvias para los angloparlantes competentes, mientras que el análisis propuesto por el reduccionista sobre el "bien" no es intuitivamente obvio. Sabemos cuándo una afirmación es contradictoria, y "El placer no es bueno" no suena contradictorio. Por lo tanto, el reduccionista no está en

condiciones de apelar simplemente a su teoría para justificar el rechazo de las premisas que Moore utilizaría contra ella. El reduccionista tampoco debería cuestionar la fiabilidad general de nuestros juicios de consistencia e inconsistencia. Pues este tipo de juicios son la esencia de nuestro conocimiento de la lógica, sin el cual ninguno de nosotros está en condiciones de avanzar o evaluar ningún argumento.

La segunda respuesta a Moore comienza con la observación de que incluso para las personas que entienden una palabra determinada, a menudo es difícil identificar de antemano la definición correcta de esa palabra. Por ejemplo, una persona que conoce la palabra "círculo" puede tener dificultades para encontrar la definición "el conjunto de todos los puntos que están a una distancia determinada de un punto determinado". Así, a menudo una pregunta puede parecerle a alguien abierta aunque no lo sea, simplemente porque la persona no se da cuenta de cuál es el análisis correcto de su propio concepto. Tal vez la pregunta de Moore, "¿Es bueno hacer lo que aumenta la cantidad de disfrute en el mundo?", sólo nos parece abierta porque al principio no nos damos cuenta de cuál es el análisis correcto de "bueno".[10]

Esta respuesta se basa en una lectura poco caritativa de Moore. Moore no tiene por qué suponer que el análisis correcto de un determinado concepto es siempre inmediatamente obvio para cualquiera que lo tenga. En lugar de ello, Moore puede arreglárselas sólo con la suposición de que, una vez que se ha enunciado una definición específica, y una vez que los hablantes competentes de la lengua han comprendido a fondo y reflexionado cuidadosamente sobre esa definición, sus intuiciones lingüísticas proporcionan al menos una guía *prima facie* de la corrección de la definición. Esta suposición es mucho más débil y razonable que la de que el análisis de los conceptos propios es siempre inmediatamente obvio, y es difícil ver cómo cualquiera que

se dedique a intentar definir palabras o conceptos -como el reduccionista analítico- puede rechazarla. Porque sin ella, es difícil ver cómo podríamos saber cuál es la definición de algo.

Las intuiciones a las que Moore apela -por ejemplo, que "El placer no es bueno" es consistente, y que "¿Es bueno el placer?" no significa "¿Es el placer conducente al placer?"[11]- son las que se producen después de que los hablantes competentes de inglés hayan comprendido a fondo la definición de "bueno" que se discute y hayan reflexionado cuidadosamente sobre ella. Por lo tanto, estas intuiciones proporcionan al menos una prueba *prima facie* de la corrección de la definición. Es decir, a falta de motivos específicos e independientes para rechazar estas intuiciones o considerarlas poco fiables, proporcionan una razón adecuada para rechazar la definición de los reduccionistas. Dado que los reduccionistas no han proporcionado tales motivos, deberíamos rechazar la definición propuesta.[12]

La pretensión de Moore de refutar todas las formas de reduccionismo analítico se basa en su capacidad de producir intuiciones similares para cada definición propuesta de "bien". Consideremos otras posibles definiciones y algunas premisas que podrían utilizarse en los argumentos mooreanos contra ellas:

Análisis propuesto
x es bueno = Dios aprueba *x*.

Premisa mooreana
"Dios es bueno" no significa "Dios se aprueba a sí mismo".

x es correcto = *x* está de acuerdo con las costumbres sociales imperantes.
"¿Es correcto obedecer las costumbres sociales?" no significa "¿Obedecer las costumbres sociales está de acuerdo con las

costumbres sociales?"

x es bueno = apruebo *x*.
"¿Son realmente buenas las cosas que apruebo?" no significa "¿Apruebo realmente las cosas que apruebo?"

x es bueno = *x* promueve la vida.[13]
"La vida es buena" no significa "¿La vida se promueve a sí misma?"

No puedo discutir aquí todos los posibles análisis de "bien". Pero los ejemplos anteriores (incluyendo el caso de la definición que implica el disfrute) sugieren que Moore podría, de hecho, aducir premisas relevantes y plausibles para su argumento en contra de cualquier forma de reduccionismo analítico. No estoy haciendo un argumento inductivo basado en cinco casos. Más bien, estoy sugiriendo que al considerar estos cinco casos, se puede ver un patrón, un método que Moore podría utilizar para generar argumentos de tipo Pregunta Abierta en otros casos.

4.3 La Guillotina de Hume

4.3.1 La ley de Hume: una declaración inicial

Estrechamente relacionada con la refutación de Moore del reduccionismo analítico hay otra doctrina en metaética conocida como la del "Problema del ser y el deber ser", "La Guillotina de Hume", o "Ley de Hume". Se trata de la idea de que es imposible derivar de forma válida una conclusión evaluativa a partir de premisas totalmente no evaluativas o, como se dice a veces, no se pueden deducir conclusiones sobre lo que debe ser a partir de premisas sobre lo que es. El *locus classicus* de esta idea es el siguiente pasaje muy trillado de David Hume:

> *En todos los sistemas de moralidad que he encontrado hasta ahora, siempre he observado que el autor procede durante algún tiempo en*

> *la forma ordinaria de razonamiento, y establece el ser de un Dios, o hace observaciones relativas a los asuntos humanos; cuando de repente me sorprende encontrar, que en lugar de las cópulas habituales de proposiciones, es, y no es, no encuentro ninguna proposición que no esté conectada con un debe, o un no debe. Este cambio es imperceptible; pero es, sin embargo, de la última consecuencia. Porque como este debe o no debe expresa alguna nueva relación o afirmación, es necesario que sea observada y explicada; y al mismo tiempo que se dé una razón de lo que parece totalmente inconcebible, cómo esta nueva relación puede ser una deducción de otras, que son completamente diferentes de ella.*[14]

Nótese que Hume no se limita a observar que de "*x* es el caso" no se puede inferir válidamente "*x* debe ser el caso". Más bien, está haciendo la observación general de que ningún conjunto de hechos sobre lo que es el caso implica ninguna afirmación sobre lo que debería ser. Otra forma de enunciar el principio es decir que cualquier descripción (no evaluativa) del mundo puede combinarse de forma coherente con cualquier sistema de valores~ o incluso con el nihilismo.[15]

Consideremos un simple desafío a la Ley de Hume. Supongamos que yo declaro que el comunismo es malo. Usted me pregunta por qué lo creo. Yo digo: "Ha causado decenas de millones de muertes, pobreza generalizada y mucho sufrimiento humano. Es razonable creer que seguirá teniendo esos efectos en el futuro. Por lo tanto, es un mal sistema de gobierno". Supongamos que mi afirmación histórica está adecuadamente establecida, y supongamos que usted acepta mi generalización de que cabe esperar que el comunismo siga teniendo efectos similares. ¿No está justificada mi conclusión? Seguramente no se trata de una simple falacia lógica que he cometido. Sin embargo, las premisas de mi argumento son totalmente no evaluativas, mientras que la conclusión es evaluativa.

Según la doctrina estándar, que yo comparto, el argumento no salva la brecha del es al debe porque contiene ciertas premisas implícitas que son evaluativas. Específicamente, el argumento asume implícitamente que la vida es buena y que la pobreza y el sufrimiento humano son malos -si no asumimos estas cosas, entonces la conclusión de que el comunismo es malo no se sigue. Ahora bien, estos supuestos pueden ser perfectamente razonables, pero son, no obstante, premisas evaluativas; por tanto, el argumento en cuestión no es un ejemplo de derivación de una conclusión evaluativa a partir de premisas totalmente no evaluativas.[16] Un punto similar se aplica a todas las formas familiares en las que típicamente argumentamos las conclusiones evaluativas; generalmente argumentamos que *x* es bueno presuponiendo algunos estándares evaluativos. Si sabemos que *y* es bueno, entonces podemos argumentar que *x* es bueno mostrando que *x* conduce a *y*. Si sabemos que la gente debe hacer *A* cuando está en circunstancias del tipo *C*, entonces podemos argumentar que *S* debe hacer *A* mostrando que S está en *C*.

Esto demuestra que algunos argumentos morales conocidos no logran salvar la brecha entre el es y el debe. Pero, ¿qué justifica la conclusión general de que ningún argumento puede hacerlo? Creo que la justificación inicial de la Ley de Hume es intuitiva: simplemente es muy difícil ver, a primera vista, cómo una mera descripción de cómo es el mundo de hecho (en aspectos no evaluativos) podría descartar algún sistema de valores que de otro modo sería posible. Es igualmente difícil ver cómo un nihilista que reconociera las propiedades naturales de las cosas pero no valorara nada podría ser condenado por contradecirse a sí mismo; parece que el nihilismo seguiría siendo siempre una posición coherente frente a cualquier hecho puramente natural. No digo que sería una posición plausible o justificada, pero sería coherente.

Otra razón que los filósofos han dado para la existencia de la Ley de

Hume es que es sólo un caso especial de un principio general de la lógica: no se puede deducir válidamente una conclusión que aplique un predicado *P*, a partir de premisas que no contengan ese predicado en ninguna parte.[17] Esto no es nada especial sobre los predicados evaluativos; se aplica a cualquier predicado. Por ejemplo, no se puede inferir una conclusión que afirme que algo es rojo a partir de premisas que no contengan en ninguna parte la palabra "rojo".

Antes de pasar a las objeciones a la Ley de Hume, es importante comprender cuál es, en mi opinión, el significado de la doctrina. No afirmo que la Ley de Hume refute el reduccionismo ético.[18] Más bien, veo la Ley de Hume como un lema importante en mi argumento más amplio de que no se puede dar cuenta del conocimiento moral sin apelar a algo como la intuición ética. Para ello, sostengo: (a) que no podemos conocer las verdades morales por observación, (b) que no podemos conocer las verdades morales (no triviales) deduciéndolas de verdades no morales, (c) que no podemos conocer las verdades morales (no triviales) por análisis conceptual, y (d) que no podemos conocer las verdades morales por razonamiento científico o inferencia a la mejor explicación. Esta interpretación del uso que hago de la Ley de Hume tiene implicaciones sobre cómo debemos tratar las supuestas excepciones a la Ley: tales excepciones sólo son importantes si representan posibles formas de llegar a conocer verdades morales sobre la base de verdades no morales. Así, si se demostrara que existe una inferencia deductiva válida que contiene sólo premisas no evaluativas y una conclusión evaluativa, pero por alguna razón esta inferencia no pudiera ser la fuente de nuestro conocimiento de la conclusión, entonces sería irrelevante para mis propósitos aquí. Sería una excepción a la Ley de Hume tal como se enuncia tradicionalmente, pero no una que perjudique mi argumento en este capítulo.

*4.3.2 El desafío de Searle

John Searle nos pide que consideremos la siguiente secuencia de afirmaciones:

1. Jones pronunció las palabras: "Por la presente prometo pagarte, Smith, cinco dólares".
2. Jones prometió pagar a Smith cinco dólares.
3. Jones se obligó a pagar a Smith cinco dólares.
4. Jones tiene la obligación de pagar a Smith cinco dólares.
5. Jones debe pagar a Smith cinco dólares.

(1) por sí mismo no implica (2); para concluir que Jones prometió pagar a Smith cinco dólares, debemos tener un conjunto de conocimientos previos sobre las convenciones de la sociedad de Jones (las que determinan los significados de las palabras de Jones, cómo se hacen las promesas, etc.) y las circunstancias de la expresión de Jones (Jones no estaba bromeando, estaba leyendo las líneas para una obra de teatro, etc.). Pero esta información adicional se compone enteramente de hechos naturales. Así que imaginemos que toda esta información se añade a las premisas del argumento, y preguntemos si el argumento entonces salvaría la Ley de Hume.

(3) se deduce de (2), dice Searle, porque forma parte del concepto de promesa el hecho de que una promesa es el compromiso de una obligación; esto es lo que distingue la promesa de hacer algo de la mera predicción de que uno lo hará o la expresión de la intención de hacerlo.

De (3) se deduce, al menos, que en igualdad de condiciones (siempre que no ocurra nada más que elimine la obligación), Jones tiene la obligación de pagar a Smith cinco dólares.

De (4) se deduce que, en igualdad de condiciones (siempre que no exista una obligación más importante y conflictiva), Jones debe pagar a Smith cinco dólares. Por lo tanto, parece que al menos podemos obtener una conclusión como "En igualdad de condiciones, Jones debe pagar a Smith cinco dólares" a partir de premisas no evaluativas.[19]

¿Qué hay de malo en este argumento? Para empezar, vamos a distinguir dos sentidos de "emprender", que llamaré el sentido fuerte y el sentido débil. En el sentido fuerte, asumir una obligación es colocarse con éxito bajo una obligación, de modo que se sigue lógicamente que uno está de hecho bajo una obligación. En el sentido débil, contraer una obligación es, a grandes rasgos, pretender colocarse bajo una obligación y/o estar de acuerdo en aceptar dicha obligación. (El análisis exacto no es importante aquí.) Una persona que asume una obligación en el sentido débil no tiene por qué tener una obligación de hecho. Consideremos una analogía: supongamos que se llama a cobro revertido, y que yo acepto los cargos, pero debido a un error, la compañía telefónica nunca me cobra. El hecho de que yo haya "aceptado" los cargos no implica que los reciba realmente. En un sentido similar, una persona podría "aceptar" una obligación sin llegar a tenerla. ¿Por qué puede ocurrir esto? Supongamos, por ejemplo, que el nihilismo es cierto: que no hay propiedades morales en absoluto. En ese caso, aunque las personas "acepten" y "asuman" (en sentido débil) obligaciones, sus compromisos siempre fracasarían a la hora de ponerlos realmente bajo obligaciones.

Por supuesto, creo que el nihilismo es falso y que sí tenemos obligaciones de cumplir nuestras promesas. Pero Searle no ha refutado el nihilismo y, por tanto, no ha demostrado un estado de "deber". No hay ningún hecho puramente natural que el nihilista sea incapaz de reconocer. El nihilista, confrontado con la derivación de Searle, debería conceder que Jones hizo una promesa, cuando esto se entiende como la aceptación

o el compromiso de una obligación en el sentido débil; pero debería negar que Jones asumió una obligación en el sentido fuerte de colocarse exitosamente bajo una obligación.

Searle podría argumentar que la noción débil de asumir obligaciones no hace justicia a nuestro concepto de "prometer"; si es así, el nihilista debería decir, entonces nadie promete realmente; como mucho pretendemos prometer. Las convenciones de nuestra sociedad dictan que decir palabras como "prometo pagarle cinco dólares" se considera una obligación, pero estas convenciones no pueden hacer que cosas como las "obligaciones" existan realmente; no podemos hacer que el nihilismo sea falso por decreto.[20]

*4.3.3 El desafío de Geach

Peter Geach propone el siguiente supuesto contraejemplo a la Ley de Hume:

1. Si Evan prometiera adoptar alguna práctica, entonces la adoptaría. (Premisa.)
2. Si Evan pronunciara la frase *W*, estaría prometiendo adoptar la práctica de hacer algo malo dos veces al día. (Premisa.)
3. Nadie debería adoptar la práctica de hacer algo malo dos veces al día. (Preludio.)
4. Si Evan pronunciara *W*, adoptaría la práctica de hacer algo malo dos veces al día. (De 1, 2.)
5. Por lo tanto, Evan no debería pronunciar *W*. (De 3, 4.)[21]

Imaginamos que *W* es una frase como "Prometo hacer dos cosas malas

cada día", y que Evan es una persona con un fuerte compromiso con el cumplimiento de las promesas, de modo que (1) es verdadera. Lo anterior parece ser una inferencia válida desde las premisas (1), (2) y (3) a la conclusión (5). (5) es claramente evaluativa. (1) y (2) parecen ser no evaluativas. Si (3) fuera no evaluativa, el contraejemplo parecería estar en orden, pero de hecho (3) es evaluativa. Sin embargo, Geach observa que (3) es una verdad analítica. Por definición, nadie debe hacer lo que está mal; a fortiori, nadie debe hacer algo malo dos veces al día. Así que, al menos, Geach parece haber mostrado cómo pasar de afirmaciones no evaluativas y/o tautologías a una conclusión evaluativa no trivial.

El ejemplo de Geach tiene al menos dos defectos.[22] En primer lugar, el argumento no es válido. Asumo que la inferencia de (3) y (4) a (5) es válida sólo si lo siguiente es una forma válida de inferencia:

6. *S* no debe hacer *A*.

7. Si *S* hiciera *B*, *S* haría *A*.

8. Por lo tanto, *S* no debería hacer *B*.

Para ver el problema de esto, imagine el siguiente escenario. Jon es un juez que está a punto de dictar sentencia contra María, una traficante de marihuana condenada. El delito de María es menor en el peor de los casos; sin embargo, Juan siente un odio intenso e irracional hacia todos los consumidores de drogas, por lo que está decidido a condenar a María a cadena perpetua o a la muerte. Podría condenar a María a una breve pena de prisión, pero no lo haría. Considere ahora la siguiente inferencia:

9. Juan no debería condenar a María a cadena perpetua.

10. Si Juan se abstuviera de condenar a María a la muerte, entonces condenaría a María a cadena perpetua.

11. Por lo tanto, Juan no debería abstenerse de condenar a María a

muerte (es decir, debería condenar a María a muerte).

(9) es cierto porque lo que Juan debería hacer es condenar a María a una breve pena de prisión como máximo. (10) es verdadera por lo que hemos estipulado sobre la psicología de Juan. Pero (11) es falsa. En general, esta forma de inferencia es inválida porque es posible que haya una situación en la que lo que uno debería hacer es hacer *B* y abstenerse de hacer *A*, pero lo que uno haría realmente después de hacer B es hacer también *A*. En tal situación, (6) y (7) son verdaderas y (8) es falsa.

El segundo problema del ejemplo de Geach es que la premisa "Si Evan prometiera adoptar alguna práctica, entonces la adoptaría" es evaluativa. La lectura pertinente de (1) es:

$(\forall x)$(Evan promete hacer x ⇨ Evan hace x)

("para todo x, si Evan prometiera hacer x, entonces Evan haría x"). Esto es evaluativo porque la siguiente instancia de sustitución es evaluativa:

Evan promete actuar mal ⇨ Evan actúa mal.

Esto es evaluativo porque, una vez que fijamos los hechos naturales sobre cómo se comportaría Evan al hacer tal promesa, el que toda la frase sea verdadera depende entonces de si ese comportamiento sería incorrecto. Por ejemplo, supongamos que, si Evan prometiera actuar de forma incorrecta, posteriormente (pensando que estas cosas son incorrectas) robaría un caramelo a su sobrina de cinco años y pediría una hamburguesa en McDonald's. Estos son los hechos naturales de cómo se comportaría Evan. Que "Evan promete actuar mal ⇨ Evan actúa de forma errónea" sea cierto depende entonces de si robar el caramelo o pedir la hamburguesa serían de hecho erróneos. Si lo fueran, entonces la frase citada es verdadera; si no lo fueran (y ninguna otra cosa que Evan hiciera estaría mal), entonces la frase citada es falsa. Por eso "Evan promete actuar mal ⇨ Evan actúa mal" es evaluativa. Se trata de una

aplicación del criterio para los enunciados evaluativos desarrollado más adelante en la sección 4.3.5.

*4.3.4 Impugnación de Prior

A. N. Prior ofrece la siguiente inferencia que supuestamente abarca la brecha es-debería:

(A)

1. Beber té es habitual en Inglaterra.
2. Por lo tanto, beber té es común en Inglaterra, o todos los neozelandeses deberían ser fusilados.[23]

La inferencia es perfectamente válida, "debería" aparece en la conclusión, pero no aparece en la premisa.

Podríamos responder negando que la conclusión sea genuinamente evaluativa. Pero esto sólo abriría la posibilidad de una inferencia diferente que salvara la brecha entre el es y el debe:

(A")

1". Beber té no es habitual en Inglaterra.
2. Beber té es común en Inglaterra, o todos los neozelandeses deberían ser fusilados.
3. Por lo tanto, todos los neozelandeses deberían ser fusilados.

Si (2) es evaluativo, entonces el argumento (A) abarca la brecha es-debería; si (2) es no evaluativo, entonces (A") abarca la brecha es-pensamiento.

Sin embargo, este punto tiene poca relación con el problema del conocimiento moral, ya que ninguna de estas inferencias proporciona un modelo general plausible de conocimiento moral. Aunque (A) puede

darnos conocimiento de una "conclusión evaluativa" en algún sentido, no es una conclusión evaluativa interesante. No es una conclusión que, por ejemplo, podamos utilizar para guiar nuestra acción, ni implica que algo tenga realmente alguna propiedad evaluativa. La conclusión de (A"), por otra parte, es una evaluación interesante, pero (A") no proporciona ningún modelo para el conocimiento moral, ya que no hay ninguna manera obvia de que podamos conocer sus dos premisas. Si (1") fuera cierta, podríamos descubrirla fácilmente por medios empíricos, pero no hay forma aparente de descubrir simultáneamente que (2) es cierta sin basarnos en un supuesto evaluativo.

Lo que sería útil sería un argumento que combinara las virtudes de (A) y (A"): un argumento que partiera de algo obviamente verificable empíricamente, como (1), y concluyera con una evaluación interesante, orientadora de la acción, como (3). Pero el ejemplo de Prior no sugiere ninguna manera de construir tal argumento. Obviamente, no podemos simplemente combinar las premisas de (A) y (A"): eso nos daría un argumento válido de premisas no evaluativas a una conclusión evaluativa. Pero como las premisas serían inconsistentes, este argumento no proporcionaría ningún modelo plausible de cómo llegamos a conocer las verdades morales.

*4.3.5 La prueba de Karmo

La respuesta anterior a Prior puede dejarnos preocupados: una vez que se ha mostrado que la barrera es-debería no es en principio inexpugnable, ¿cómo podemos estar seguros de que todo argumento que penetre esa barrera resultará tan inútil para dar cuenta del conocimiento moral como el ejemplo de Prior? ¿Podemos idear una prueba general para alguna forma útil de la Ley de Hume?

Toomas Karmo ha proporcionado un argumento de este tipo.[24]

Supongamos que hay una clase de enunciados incontrovertiblemente evaluativos que atribuyen claramente propiedades evaluativas a las cosas -como "Todos los neozelandeses deberían ser fusilados" y "El placer es bueno"- y una clase de enunciados incontrovertiblemente descriptivos que atribuyen claramente sólo propiedades no evaluativas -como "Beber té es común en Inglaterra". También hay enunciados cuyo estatus no está claro en un principio, como "Beber té es común en Inglaterra, o todos los neozelandeses deberían ser fusilados" y "Todo lo que dice Alfie es cierto". Según Karmo, estas últimas afirmaciones pueden ser evaluativas o no evaluativas, dependiendo de cómo sea el mundo en aspectos no evaluativos. Por ejemplo, si beber té es de hecho común en Inglaterra, entonces "Beber té es común en Inglaterra, o todos los neozelandeses deberían ser fusilados" es no evaluativo (aproximadamente, porque su verdad no depende de si los neozelandeses deberían ser fusilados); de lo contrario, la afirmación es evaluativa. Del mismo modo, si Alfie ha hecho de hecho algunas afirmaciones evaluativas y no ha hecho afirmaciones descriptivas falsas, entonces "Todo lo que Alfie dice es verdad" es evaluativo, ya que su verdad depende entonces de si las afirmaciones evaluativas de Alfie eran verdaderas; de lo contrario, es no evaluativo.

A continuación se presenta un enunciado más preciso de la idea de Karmo. Sea un "sistema de valores" una conjunción consistente máxima de enunciados evaluativos incontrovertibles (es decir, una conjunción de enunciados evaluativos incontrovertibles que es consistente y tal que no se pueden añadir más enunciados de este tipo sin crear una inconsistencia). Del mismo modo, dejemos que un "mundo posible" sea una conjunción máxima consistente de enunciados incontrovertiblemente descriptivos. El "mundo real" es el mundo posible que es verdadero (la descripción correcta del mundo). Asumo el siguiente principio, que llamaré de Determinación:

Determinación Un mundo posible junto con un sistema de valores son consistentes con a lo sumo una asignación de valores de verdad para todos los enunciados.

Es decir, si *w* es un mundo posible, v es un sistema de valores, y 5 es un enunciado cualquiera, entonces (*w* & *v*) implica 5, o (*w* & *v*) implica -*S*.[25]

Ahora, para los enunciados cuyo estatus como evaluativo o descriptivo no está inicialmente claro, clasifíquelos como sigue: un enunciado 5 es evaluativo en un mundo posible *w* si, dado *w*, el valor de verdad de 5 depende de qué sistema de valores es correcto. En otras palabras, 5 es evaluativo en *w* siempre que haya dos sistemas de valores tales que uno de ellos, combinado con los enunciados incontrovertiblemente descriptivos en *w*, implique que 5 es verdadero, mientras que el otro, combinado con los mismos enunciados descriptivos, implique que 5 es falso. Esta caracterización da como resultado, por ejemplo, que "Todo lo que dice Alfie es verdadero" es evaluativo en cualquier mundo posible en el que Alfie haya hecho un enunciado evaluativo y no haya hecho ningún enunciado descriptivo falso, pero no evaluativo en cualquier otro mundo posible.

Lo que nos gustaría ver es si puede haber un argumento tal que:

i) Sus premisas sean realmente verdaderas,

ii) Sus premisas no son evaluables en el mundo real,

iii) Sus premisas implican su conclusión,

iv) Su conclusión es evaluativa en el mundo real, y

v) Su conclusión no es trivial.

Las condiciones (*ii*), (*iii*), y (*iv*) son necesarias para que el argumento cuente como que abarca la brecha es-pensamiento. La condición (*i*) es

necesaria para que el argumento cuente como una forma de obtener conocimiento. La condición (*v*) es necesaria para que el argumento sea interesante en relación con el problema del conocimiento moral: la conclusión no debe ser una mera tautología, como "Los actos correctos deben realizarse" o "Las cosas malas son malas" (estas afirmaciones se deducen trivialmente de cualquier conjunto de premisas). Asumo que un enunciado evaluativo es no trivial sólo si hay al menos un sistema de valores que lo rechaza.

Supongamos ahora para la *reductio* que existe un argumento que satisface las condiciones (*i*)-(*v*). Sean P_1, P_2, ... las premisas de este argumento, y sea *C* su conclusión. Sea *w* el mundo real. Entonces tenemos lo siguiente:

1. P_1, P_2, ... conlleva *C*. (Condición (*iii*).)
2. C es no trivial y evaluativo en *w*. (Condiciones (*iv*) y (*v*).)
3. Existen sistemas de valores v_1 y v_2 tales que (*w* & v_1) conllevan *C* y (*w* & v_2) conllevan ~C. (A partir de 2.)
4. ~*C* implica que al menos uno de {P_1, P_2, ...} es falso. (De 1.)
5. (*w* & v_2) implica que al menos una de {*p*), P_2, ... } es falsa. (De 3, 4.)
6. Hay una premisa específica en {P_1, P_2, ... }, llámese P_i, tal que (*w* & v_2) implica ~P_i. (De 5 y de la Determinación.)
7. *w*, junto con el sistema de valores correcto, implica P_i. (De la condición (*i*) y de la Determinación.)
8. P_i es evaluativo en *w*. (De 6, 7.)

La conclusión (8) se desprende de (6) y (7) porque, dado *w*, la verdad de P_i depende de cuál sea el sistema de valores correcto: P_i es verdadera según el sistema de valores correcto pero sería falsa según v_2. Pero (8)

contradice la condición (*ii*). Por tanto, no hay ningún argumento que satisfaga todas las condiciones (*i*)-(*v*).

Considero que la prueba de Karmo tiene éxito. Sin embargo, Stephen Maitzen ha objetado el argumento.[26] Sin entrar en detalles, observo que la objeción de Maitzen se basa en una crítica a la taxonomía de Karmo, una crítica que depende de la suposición de que el nihilismo -definido como la tesis de que todo estado evaluativo es falso- es lógicamente posible. Entre otras cosas, Maitzen argumenta que la taxonomía de Karmo es defectuosa ya que implica que el siguiente enunciado es evaluativo en todo mundo posible:

B. Todo enunciado evaluativo es falso, o la tortura de bebés está mal.

Según la taxonomía de Karmo, esto es evaluativo en todos los mundos posibles, ya que siempre hay un sistema de valores que hace que la afirmación sea verdadera (cualquier sistema de valores que prohíbe la tortura de bebés hace que (B) sea verdadera) y otro sistema de valores que hace que la afirmación sea falsa (cualquier sistema de valores que aprueba la tortura de bebés hará que ambos disyuntos de (B) sean falsos). Pero Maitzen dice que el enunciado debe ser clasificado como no valorativo en al menos algunos mundos posibles, ya que:

1. El nihilismo es lógicamente posible. (Premisa.)
2. El nihilismo implica (B). (Por V-introducción.)
3. Por lo tanto, es posible que tanto el nihilismo como (B) se mantengan. (De 1, 2.)
4. En cualquier mundo posible en el que el nihilismo y (B) sean válidos, (B) no es evaluativo. (En tal mundo, toda afirmación evaluativa es falsa, ya que el nihilismo se mantiene. Pero (B) también se sostiene, por lo que (B) no debe ser un enunciado evaluativoInent).

5. Así que hay al menos un mundo posible en el que (B) no es evaluativo. (De 3, 4.)

Este ejemplo, dice Maitzen, "falsifica inmediatamente la taxonomía de Karmo de las oraciones éticas".[27]

Maitzen se equivoca porque el nihilismo -entendido como la tesis de que todo "enunciado valorativo" en el sentido de Karma es falso- es lógicamente imposible. Una razón para ello es que el propio nihilismo parece ser una afirmación evaluativa;[28] por lo tanto, al negar todas las afirmaciones evaluativas, se contradice a sí mismo. Otra razón es que, en cualquier mundo posible, tanto "Torturar bebés está mal" como "No es cierto que torturar bebés esté mal" son enunciados evaluativos en el sentido de Karmo. Pero es lógicamente necesario que al menos una de estas dos afirmaciones sea verdadera. Por tanto, es lógicamente imposible que todas las afirmaciones evaluativas sean falsas. Por supuesto, el nihilismo en algún otro sentido podría ser lógicamente posible; por ejemplo, tal vez sea posible que nada tenga una propiedad evaluativa. Pero esto no refutaría el argumento de Karmo. Del mismo modo, podría ser que el sentido de Karmo de "enunciado evaluativo" difiera del uso establecido -ciertamente difiere del uso de Maitzen-, pero esto tampoco mostraría que el argumento de Karmo no es sólido. En general, no se trata de "falsificar" la caracterización de Karmo de los enunciados valorativos, siempre que esa caracterización se entienda -como creo que debería ser- como una definición estipulativa.

Dado que la definición es puramente estipulativa, uno puede preguntarse legítimamente si la prueba de Karmo es relevante para lo que nos interesaba inicialmente -en particular, ¿ayuda la versión de Karmo de la brecha es-debería a establecer que hay un problema en la contabilidad del conocimiento moral? Creo que sí. Karmo muestra que no hay forma de conocer los enunciados evaluativos por deducción de

los enunciados no evaluativos, en su sentido de "enunciados evaluativos". Pero el sentido de Karmo incluye todos los enunciados tradicionalmente considerados "evaluativos" por los filósofos morales; por tanto, si es difícil conocer enunciados evaluativos en el sentido de Karmo, también es difícil conocer enunciados evaluativos en el sentido tradicional. Tampoco parece que Karmo haya ampliado tanto el sentido de "enunciado evaluativo" como para que podamos conocer algunos enunciados evaluativos, por ejemplo, mediante la observación: en el sentido de Karmo, un enunciado evaluativo es aquel cuya verdad, una vez fijados todos los hechos descriptivos, sigue dependiendo del sistema de valores que sea correcto. Salvo que se recurra a cosas como la intuición moral, sigue siendo difícil, a primera vista, ver cómo podemos conocer tales afirmaciones de forma no inferencial. Dado que Karmo ha demostrado que, en general, tampoco podemos conocer tales afirmaciones por deducción de afirmaciones no evaluativas, se plantea un interesante problema de conocimiento moral.

En la siguiente sección, consideramos los intentos de mostrar que podemos conocer verdades evaluativas mediante inferencias no deductivas a partir de verdades no evaluativas.

4.4 Reduccionismo sintético

Los reduccionistas sintéticos parecen escapar al argumento de la pregunta abierta,[29] sostienen que hay algún enunciado verdadero de la forma: "Para que una cosa sea buena es . . . ", donde la elipsis se rellena con términos no valorativos. Pero admiten que, como demostró Moore, lo que rellena la elipsis no será sinónimo de "para que sea bueno". Consideremos algunos casos análogos:

> Que una cosa esté caliente es que sus moléculas tengan una energía

cinética media alta.

Que una sustancia sea agua es que esté formada por moléculas de H_2O.

Para que una cosa emita un sonido es necesario que provoque una onda de compresión en el medio circundante.

Dejando a un lado las objeciones, todas estas afirmaciones son ciertas. Pero no pueden verificarse únicamente considerando los significados de las palabras "caliente", "agua" y "sonido" en inglés. La gente conoció esas palabras y las utilizó en el mismo sentido que nosotros lo hacemos actualmente durante siglos sin conocer los hechos expresados por las afirmaciones anteriores. Se necesitaban observaciones muy especializadas y razonamientos científicos para confirmar esos hechos. Del mismo modo, dice el reduccionista sintético, algún día encontraremos (si no lo hemos hecho ya) una teoría de lo que significa que algo sea bueno, y esa teoría no se conocerá simplemente sobre la base del significado de la palabra "bueno" en inglés. Por el contrario, se conocerá de forma análoga a como se conoce la veracidad de los enunciados anteriores, mediante una especie de inferencia a la mejor explicación.[30] En esta sección, sostengo que el relato de los reduccionistas sintéticos sobre el conocimiento moral fracasa: ninguna teoría sobre la naturaleza de la bondad puede conocerse de manera análoga a como se conocen las teorías anteriores sobre el calor, el agua y el sonido. Al final, el reduccionista sintético tendrá que apelar a la intuición ética, al igual que los intuicionistas. Esto priva a su posición de una de sus supuestas ventajas centrales.

4.4.1 ¿Se pueden conocer los hechos morales mediante la observación?

La analogía que acabamos de establecer entre el reduccionismo ético y

diversas teorías científicas desempeña un papel bastante importante en la motivación de la reducción sintética.[1] Pero la analogía falla en un punto crucial.

Supongamos que nos dirigimos a un reduccionista sintético que sostiene que para que una cosa sea buena es que esta sea *N*, donde *N* es alguna propiedad natural, o para abreviar:

Bueno = *N*.

El reduccionista dice que esto es análogo a las teorías antes mencionadas, que abreviaré de la siguiente manera:

Calor = Energía cinética molecular.

Agua = H_2O.

Sonido = Ondas de compresión.

En cada uno de estos casos, tenemos dos conceptos diferentes que se refieren al mismo fenómeno (el concepto "agua" es diferente del concepto "H_2O", pero hay una cosa en la realidad que ambos conceptos recogen). En cada caso, el concepto de la izquierda es observacional (podemos aplicarlo a una cosa basándonos en la observación), mientras que el concepto de la derecha es teórico (lo aplicamos a una cosa basándonos en la inferencia, incluida la inferencia a una teoría científica) y se toma para explicar la naturaleza del fenómeno en cuestión. Así, inicialmente sabíamos que las cosas estaban calientes porque podíamos sentir el calor con las manos. Sabemos que las cosas tienen energía cinética molecular a partir de inferencias complejas de los datos observacionales a la teoría molecular de la materia, etcétera. Se cree que la idea de la energía cinética molecular explica la naturaleza subyacente del calor. Del mismo modo, sabemos que hay agua por observación casual, pero sabemos que hay H_2O por un razonamiento científico

complejo. Y el "H_2O" explica la naturaleza subyacente del agua.

Ahora bien, en la analogía del reduccionista, ¿a qué término se supone que es análogo "bueno": al término de la izquierda del "=" o al de la derecha? ¿Se supone que "bueno" es como "calor" o como "energía cinética molecular"? Inicialmente, parece que el "bien" se supone que es análogo al "calor", porque la teoría reduccionista se supone que explica la naturaleza subyacente de la bondad en términos de propiedades naturales, al igual que la teoría científica explica la naturaleza subyacente del calor en términos de movimientos de las moléculas. Ciertamente, los reduccionistas éticos no se consideran a sí mismos explicando la naturaleza subyacente de algunas propiedades naturales en términos de propiedades morales.

Esto sugiere que los conceptos morales deberían ser observacionales, al igual que el "calor", el "agua" y el "sonido". La razón por la que creemos, por ejemplo, que el agua = H_2O es, a grandes rasgos, que (*i*) tenemos conciencia directa e independiente (es decir, preteórica) de la presencia del agua, y (*ii*) la teoría de que esta sustancia de la que somos conscientes está compuesta por moléculas de H_2O ayuda a explicar muchas de sus características observables (de nuevo, características que conocemos independientemente de nuestra teoría científica). Si nos tomamos en serio las analogías del reduccionista sintético, deberíamos decir que tenemos una conciencia directa e independiente de la presencia de la bondad, y que la teoría de que la bondad es idéntica a *N* ayuda a explicar muchas de las características observables de la bondad.

El problema es que, de hecho, las propiedades morales son totalmente inobservables.[31] El valor moral no se parece a nada, ni suena como nada, ni se siente (al tacto) como nada, ni huele como nada, ni sabe como nada. La única manera plausible de sostener que tenemos conciencia directa de los hechos morales sería apelar a una "intuición

ética" o a un "sentido moral", que es precisamente el tipo de cosa "misteriosa" y "espeluznante" que el reduccionista quiere evitar.

Gilbert Harman analiza un supuesto ejemplo de observación moral.[32] Ves a un grupo de matones echar gasolina a un gato y prenderle fuego. Inmediatamente piensas que lo que están haciendo está mal. Según Harman, "no necesitas averiguar nada; puedes ver que está mal". Pero esta cita es engañosa por sí misma; Harman no cree que la maldad de la acción modifique de algún modo los rayos de luz que viajan hasta sus ojos, ni que su supuesta capacidad para ver la maldad explique en última instancia su conocimiento moral. Más bien piensa que usted tiene creencias morales de fondo, o una sensibilidad moral, que hace que, al ver un acontecimiento como el descrito, inmediatamente (sin pensar en ello) se forme la creencia de que "eso está mal". Puesto que "ver" lo que está mal depende de creencias morales previas, no ayuda en la explicación última de cómo conocemos las verdades morales.

¿Por qué considera Harman que se trata de un caso de observación? Parece pensar que "observar" que *p* es simplemente formar una creencia de que *p* como resultado bastante directo de la percepción.[33] Esto es un error; una creencia formada como resultado bastante directo de la percepción no tiene por qué ser una observación. Supongamos que soy supersticioso y cada vez que veo un gato negro, creo inmediatamente (sin pensarlo) que un avión se va a estrellar al día siguiente. Ahora veo un gato negro. ¿Observo así un accidente de avión? No. Se necesita algo más para la observación. Para que alguien observe que un objeto es *F*, donde *F* es alguna propiedad, debe haber una forma en que las cosas *F* se ven (o suenan, huelen, etc.), y el objeto debe verse (sonar, oler, etc.) de esa manera.[34] Por ejemplo, tengo un par de zapatos de cuero artificial. La mayoría de la gente puede percibir al tacto que no son de cuero, porque hay una forma de sentir el cuero real, y estos zapatos no se sienten así. El

punto de interés aquí es que no hay tal cosa como la forma en que se ven las acciones ilícitas o la forma en que se ven las acciones permisibles. Por eso no se puede ver literalmente, con los ojos, que una acción es incorrecta.[35]

Consideremos otro argumento a favor de la posibilidad de la observación moral. Michael Moore argumenta que, puesto que las propiedades morales son propiedades naturales (según su postura ética naturalista), y puesto que podemos percibir las propiedades naturales relevantes, podemos por tanto percibir las propiedades morales. En resumen, Moore considera que el siguiente argumento es válido:

Puedo percibir *N*.

Bien = *N*.

Por lo tanto, puedo percibir la bondad.[36]

Antes de abordar esto, debo hacer una corrección: la cuestión relevante no es si uno puede percibir la bondad; es si uno puede percibir que una cosa particular es buena. Se trata de dos cuestiones diferentes, y sólo la segunda tiene que ver con el problema del conocimiento moral. Para ilustrarlo, comparemos este caso: Puedo ver tu coche (tu coche refleja la luz a mis ojos, etc.) sin estar en condiciones de saber que el coche es tuyo. Si pudiera ver que es su coche, eso me permitiría saber que es suyo, pero el mero hecho de ver su coche no le da ese conocimiento. Así que modifiquemos el argumento de Moore de la siguiente manera:

Puedo percibir que *x* es *N*.

Bien = *N*.

Por lo tanto, puedo percibir que *x* es bueno.

Si esta inferencia fuera válida, Moore tendría una explicación viable

de cómo podemos conocer las verdades morales. Pero, de hecho, la inferencia no es válida. Tiene la misma forma que la siguiente inferencia:

> Puedo percibir que *x* es caliente.
>
> Calor = Energía cinética molecular.
>
> Por lo tanto, puedo percibir que las moléculas de *x* tienen una energía cinética media elevada.

Aquí utilizo uno de los ejemplos de los reduccionistas en su contra. Las dos primeras premisas pueden ser ciertas, pero la conclusión es ciertamente falsa. Por ejemplo, uno puede percibir por el tacto que una estufa está caliente. Pero uno no percibe, por el mero hecho de sentirla, que la estufa está compuesta por moléculas que se mueven rápidamente (aunque eso es lo que realmente es que la estufa esté caliente). Hay que hacer una sofisticada teorización para averiguarlo. Del mismo modo, aunque el Bien = *N*, y aunque uno pueda percibir que *x* es *N*, no percibe por tanto que *x* sea bueno.[37]

Los hechos de valor, por tanto, son inobservables. Esto es un problema para las analogías del reduccionista sintético: fuimos capaces de descubrir que el calor = la energía cinética molecular, sólo porque primero pudimos identificar qué cosas estaban calientes por observación. Si no tenemos conocimiento preteórico de qué cosas son buenas, entonces no podemos descubrir que Bien = *N* de ninguna manera análoga. La situación sería como la siguiente: alguien te dice que existe una propiedad inobservable llamada "torfness". No se le da ninguna información inicial sobre la naturaleza de torfness o qué cosas en el mundo, si las hay, son torf. Sólo se le dice que la torfness es reducible a alguna propiedad científica. Su tarea: averiguar a qué propiedad científica es reducible torfness.

Parece claro que no tienes suficiente información para llevar a cabo la tarea. Del mismo modo, sin información inicial sobre la naturaleza de la bondad o sobre qué cosas la tienen, no tendríamos ninguna base para formular ninguna hipótesis sobre a qué propiedad natural podría ser reducible la bondad. Como acabamos de ver, la observación no nos proporciona esa información inicial. Tampoco, según los naturalistas, la intuición. Uno podría estar tentado a proponer que simplemente nos basemos en nuestras creencias de sentido común sobre el bien y el mal. Pero, a menos que estemos justificados en pensar que estas creencias son generalmente verdaderas, esto sería arbitrario, y todavía no hemos visto ninguna explicación de por qué esas creencias están justificadas. Además, si el naturalista pudiera dar cuenta de nuestro conocimiento inicial de la naturaleza del valor o de qué cosas son buenas, no necesitaría la historia sobre la inferencia a la mejor explicación que los naturalistas se han esforzado tanto en desarrollar (véanse las subsecciones 4.4.2-4.4.4).

4.4.2 ¿Pueden conocerse los hechos morales por inferencia a la mejor explicación?

Algunos reduccionistas han argumentado que está justificado creer en afirmaciones morales porque proporcionan las mejores explicaciones para ciertos hechos no morales de los que tenemos conciencia independiente. Por ejemplo, la mejor explicación de por qué Hitler ordenó la Solución Final es que Hitler era un depravado; sólo una persona depravada haría algo así.[38] Además, como sabemos empíricamente que Hitler ordenó la Solución Final, y nuestra justificación para creer que era un depravado es que eso explica por qué ordenó la Solución Final, nuestro conocimiento de que era un depravado también es empírico. En resumen:

1. Las afirmaciones morales ayudan a explicar los hechos observables. Por ejemplo, la depravación de Hitler ayuda a explicar por qué ordenó la Solución Final.

2. Si una afirmación ayuda a explicar hechos observables, entonces tenemos una justificación empírica para creer en esa afirmación.

3. Por lo tanto, tenemos una justificación empírica para creer en algunas afirmaciones morales, como que Hitler era un depravado.

La conclusión de (3) debe entenderse como que podemos tener una justificación puramente empírica para creer en algunas afirmaciones morales, es decir, una justificación que no depende en última instancia, ni siquiera parcialmente, de la intuición o de cualquier otra fuente de información no observacional. La mayor parte del debate en la literatura se centra en la premisa (1). Hay dos argumentos principales para (1). Los llamo "el argumento causal" y "el argumento contrafáctico".

El argumento causal sostiene que para que *x* explique a *y* es necesario que *x* sea (en parte) la causa de *y*. La orden de Hitler de llevar a cabo la Solución Final fue causada por ciertos rasgos psicológicos suyos, como el racismo extremo, el odio y la falta de respeto por la vida humana. Según los reduccionistas éticos, estos rasgos constituían la depravación de Hitler. Por lo tanto, la depravación de Hitler provocó que ordenara la Solución Final. (Este argumento emplea el principio válido de que si *A* causa *C*, y *B* es idéntico a *A*, entonces *B* causa *C*).[39]

El argumento contrafáctico sostiene que *x* explica *y* (al menos en parte) siempre que sea cierto lo siguiente: si *x* no hubiera ocurrido, *y* no habría ocurrido. Ahora bien, si Hitler no hubiera sido un depravado, no habría ordenado la Solución Final. Dado que los hechos morales son reducibles a hechos naturales, si Hitler no hubiera sido un depravado,

tendría que haber tenido rasgos psicológicos diferentes a los que realmente tenía; por ejemplo, tendría que no haber odiado a los judíos, o haber tenido más respeto por la vida humana, etc. Y si ese fuera el caso, entonces no habría ordenado la Solución Final. Así que su depravación explica (al menos en parte) por qué ordenó la Solución Final.[40]

Estoy de acuerdo en que los hechos morales pueden ser explicativos; la premisa (1) es cierta. Pero la premisa (2) es falsa. Para ver por qué, volvamos a la analogía favorita de los reduccionistas: la teoría de que el agua = H_2O. Imaginemos que un científico intentara justificar esa teoría por los medios sugeridos por los reduccionistas éticos. Supongamos que Lavoisier hubiera publicado un artículo argumentando que está justificado creer que el agua está compuesta de hidrógeno y oxígeno, basándose en que la presencia de dicho compuesto explica ciertos hechos observables. Por ejemplo, tengo una mancha observable en mi mesa de café que fue causada por el agua. Dado que el agua = H_2O, también podemos decir que la mancha fue causada por el H_2O. Por tanto, el H_2O explica un hecho observable. Por lo tanto, tenemos una justificación empírica para creer en el H_2O.

He aquí otro argumento. Si el H_2O no se hubiera derramado en mi mesa de café, no habría una mancha de agua allí ahora. Porque recuerde, el agua se compone de H_2O, así que para que el H_2O no haya estado en mi mesa de café, tendría que no haber habido agua en mi mesa de café. Y si ese fuera el caso, no habría una mancha de agua allí ahora. Por lo tanto, la presencia (pasada) del H_2O explica un hecho actualmente observable. Así que, de nuevo, está justificado que creamos en el H_2O.

Afortunadamente, los científicos no razonan así. He aquí parte de la razón por la que los científicos creen que el agua = H_2O: si se aplica un potencial eléctrico a una muestra de agua, se forman burbujas de hidrógeno en el cátodo. En el ánodo se forman burbujas de oxígeno. La

masa de agua presente disminuye en la misma medida que aumenta la masa de hidrógeno y oxígeno presentes.[41] Además, se forman dos partes de hidrógeno por cada parte de oxígeno (en moles). Ésta es sólo una de las muchas pruebas experimentales relevantes, pero bastará para ilustrarlas. El teórico de la moral no tiene ninguna evidencia empírica que ofrecer que sea análoga a esa.

Este ejemplo sugiere que los argumentos causales y contrafácticos no muestran cómo se pueden justificar las afirmaciones morales, ya que los argumentos análogos no funcionarían para las afirmaciones científicas. Pero, ¿qué es exactamente lo que falla en esos argumentos?

Empecemos por el argumento causal. Si el reduccionismo ético es cierto, entonces las propiedades morales pueden causar efectos observables. Sin embargo, esto no explicaría cómo conocemos las afirmaciones morales. Para que conozcamos una afirmación moral sobre la base de la inferencia a la mejor explicación, la afirmación moral tendría que explicar algún hecho observable que no podría explicarse de otro modo. Por ejemplo, podemos saber que Hitler era un depravado por inferencia a la mejor explicación, si la cláusula de que Hitler era un depravado proporciona una explicación para algún hecho que no puede ser explicado sin invocar cláusulas morales. Pero, de hecho, la hipótesis de la depravación de Hitler no explica ningún hecho nuevo sobre su comportamiento que no se explique ya por afirmaciones no morales sobre él, como que odiaba a los judíos, que tenía ansias de poder, que carecía de respeto por la vida humana, etc. Como la verdad de las afirmaciones no morales de ese tipo no se discute, la explicación moral es superflua y no proporciona ninguna justificación para adoptar ninguna teoría moral. Esto no quiere decir, por supuesto, que no tengamos ninguna justificación para nuestras afirmaciones morales sobre Hitler; si ya estuviéramos justificados (quizás sobre la base de la intuición moral)

para aceptar teorías morales como que los asesinos en masa son malos, entonces podemos inferir que Hitler era malo. La cuestión es que ni nuestras teorías morales generales ni nuestros juicios morales sobre Hitler pueden establecerse únicamente sobre la base de la observación.

Lo mismo ocurre con el argumento contrafáctico. Dado que el odio racial y demás constituyen una depravación, es cierto que si Hitler no hubiera sido depravado, no habría actuado como lo hizo. Pero podemos explicar el comportamiento de Hitler igualmente bien sin decir que el odio racial y demás constituyen la depravación: si Hitler no hubiera tenido odio racial, sed de poder, etc. -independientemente de que estas cosas constituyan o no depravación- no habría actuado como lo hizo. Por lo tanto, las suposiciones morales no añaden ningún poder explicativo.

Para confirmar este diagnóstico, recordemos el ejemplo del agua/H_2O. La mancha de agua en mi mesa de café no confirma la presencia pasada de H_2O, ya que la hipótesis de que se derramó H_2O sobre la mesa no explica ningún hecho nuevo que no fuera ya explicable simplemente por la hipótesis de que se derramó agua sobre la mesa. Pero el experimento de electrólisis confirma que hay H_2O, porque la hipótesis de que hay H_2O explica de dónde proceden el hidrógeno y el oxígeno. Eso no se explica simplemente por la hipótesis de que hay agua, sin suponer que el agua está compuesta de hidrógeno y oxígeno.

*4.4.3 ¿Pueden ponerse a prueba las afirmaciones morales?

Tal vez las afirmaciones morales sean justificables porque hacen predicciones empíricas que pueden ponerse a prueba. Por ejemplo, las dos afirmaciones morales "Hitler era una persona admirable" y "Ninguna persona admirable mataría a millones de personas" conllevan la siguiente predicción: "Hitler no mataría a millones de personas". Esta predicción es comprobable y, de hecho, es falsa. Así que vemos que las

afirmaciones morales son comprobables al igual que las teorías científicas.

En caso de que usted tenga la tentación de objetar que ninguna de las afirmaciones morales implica por sí misma ninguna predicción comprobable, Nicholas Sturgeon nos recuerda que también es cierto que las afirmaciones científicas individuales no implican, en general, ninguna predicción comprobable por sí mismas; normalmente, una teoría científica sólo hace una predicción comprobable dado un amplio bagaje de otros conocimientos científicos.[42]

Incluso concediendo el último punto, no parece que esto explique el conocimiento moral. Imaginemos que Lavoisier hubiera argumentado análogamente: El conjunto de afirmaciones sobre el H_2O implica predicciones comprobables. Por ejemplo, el par de afirmaciones "Agua = H_2O" y "H_2O hierve a 212°F" juntas implican la predicción comprobable "El agua hierve a 212°F". Como esta última predicción es cierta, mis creencias sobre el H_2O se confirman.

El problema es que no todas las predicciones hechas por una teoría crean formas en las que la teoría pueda ser confirmada de forma útil. Si cada predicción se hace con la ayuda de una postura *ad hoc* separada, entonces el éxito de la teoría en la predicción será poco impresionante. Así, la teoría "Agua = H_2O" puede hacer una serie de predicciones comprobables sobre la naturaleza del agua, si le añadimos las suposiciones *ad hoc* de que el H_2O sería un líquido claro e inodoro, que herviría a 212°F, etc.; pero el éxito de estas predicciones es evidentemente nugatorio. Por el contrario, la idea de que el agua = H_2O sugiere naturalmente que el agua debería descomponerse en hidrógeno y oxígeno en una proporción de 2:1; esta predicción no requiere suposiciones ad hoc.

Todas las predicciones realizadas por las teorías morales parecen requerir posturas *ad hoc*. La teoría moral de que matar a millones de personas está mal nos permite predecir que Hitler no matará a millones de personas, sólo si hacemos la suposición inmotivada de que Hitler es moralmente admirable (o al menos no malvado). Pero la posibilidad de derivar predicciones a partir de una combinación arbitraria de afirmaciones morales no tiene ningún interés epistemológico.

Sturgeon da otro ejemplo. El par de afirmaciones "La injusticia causa inestabilidad política" y "La propiedad natural *X* constituye una injusticia" implican conjuntamente que *X* causa inestabilidad.[43] Siempre que *X* sea detectable de forma independiente, se trata de una predicción comprobable. Aquí, la suposición de Sturgeon de que la injusticia causa inestabilidad no parece ad hoc; parece una suposición natural sobre la naturaleza de la injusticia. Pero, ¿por qué? Probablemente porque asumimos (a) que la gente puede identificar la injusticia y (b) que se opone a ella. Dadas estas suposiciones, cuando la injusticia está extendida en una sociedad, es natural esperar que la sociedad sea políticamente inestable. Si las personas no tuvieran ningún medio de conocimiento moral para empezar, entonces no habría ninguna razón para esperar que las propiedades morales estuvieran correlacionadas con su comportamiento de alguna manera interesante.

El problema ahora es que nuestra capacidad de hacer una predicción basada en afirmaciones morales en este caso depende de que asumamos que la gente ya tiene alguna forma de obtener creencias morales fiables. ¿Y cuál podría ser? Si aceptamos que la gente tiene una facultad fiable de intuición moral, digamos, o que la gente puede observar hechos morales, o que puede deducir hechos morales frente a hechos no morales, entonces podríamos aceptarlo. Pero el reduccionista sintético rechaza la primera posibilidad, y ya hemos rechazado las dos últimas. Además, si el

reduccionista pudiera responder a esta pregunta, no necesitaría el presente argumento, ya tendría una explicación del conocimiento moral.

En resumen, si se nos dan todas las propiedades no morales de una cosa, la hipótesis adicional de que una de esas propiedades constituye la bondad, digamos, no parece generar naturalmente ninguna predicción nueva sobre lo que deberíamos observar, sin una postura *ad hoc* separada para cada predicción deseada. En consecuencia, partiendo de un sistema de creencias puramente no morales, no estaría justificado añadir ninguna creencia moral. Hay una excepción: parece natural y no ad hoc esperar que los seres inteligentes persigan en general el bien, se sientan gratificados cuando lo obtengan, etc. Pero esta predicción depende de la suposición previa de que las personas tienen algún medio independiente para detectar las propiedades morales, por lo que no sirve para explicar el origen del conocimiento moral.

*4.4.4 El poder unificador de las explicaciones morales

Tal vez, a pesar de que sólo proporcionan explicaciones para cosas para las que ya tenemos explicaciones no morales, las afirmaciones morales pueden seguir estando justificadas porque proporcionan mejores explicaciones que las formuladas con afirmaciones no morales. Michael S. Moore ha desarrollado esta sugerencia, argumentando que las explicaciones morales pueden lograr un tipo de unificación que no está disponible de otra manera.

Una analogía nos ayudará. Nos encontramos con que una determinada clavija no encaja en un determinado agujero. En un nivel, podríamos intentar explicar este hecho citando las trayectorias de todos los átomos del bloque y las fuerzas químicas y de otro tipo que actúan sobre cada uno de ellos. Pero una explicación más perspicaz sería: "Esto es una clavija cuadrada, y aquello es un agujero redondo". Esta última

explicación es superior por su poder unificador: dada esta explicación, sabemos qué esperar cuando intentamos meter otra clavija cuadrada en otro agujero redondo (con un diámetro menor que la diagonal del cuadrado). La explicación a nivel atómico pasa por alto esta regularidad de nivel superior: la siguiente clavija cuadrada que encontremos puede estar hecha de una sustancia diferente y seguramente tendrá una disposición diferente de los átomos, por lo que, a nivel atómico, no se le aplicará la misma explicación; tendremos que dar una explicación distinta en cada caso. Por el contrario, la explicación "clavija cuadrada, agujero redondo" identifica la característica esencial común a todas estas situaciones.[44] Ahora, supongamos que la propiedad de la bondad es idéntica a alguna disyunción de propiedades naturales, es decir:

Bueno = (*F* o *G* o *H*),

donde "*F*", "*G*" y "*H*" son términos no valorativos. Y supongamos que hay algún efecto, £, que ocurre siempre que cualquiera de *F*, G o *H* está presente. En ese caso, aunque la ocurrencia de E en una ocasión determinada podría explicarse citando la presencia de F, la mejor explicación sería la que cita la presencia de la bondad. Porque esta última explicación lograría un tipo de unificación que la primera explicación no logra: permite predecir correctamente lo que sucederá cuando ocurra otro caso de bondad.[45]

La pregunta crucial aquí es, ¿cómo se podría llegar a la idea de que el Bien = (*F* o G o *H*)? Es de suponer que el mero hecho de que cada una de varias propiedades distintas provoque un único efecto no justifica la identificación de la bondad con la disyunción de esas propiedades. Uno podría estar tentado de apelar a las creencias de sentido común sobre qué cosas son buenas, pero eso sería volver a un error que hemos criticado anteriormente. Dado que en este punto el reduccionista todavía no ha explicado con éxito cómo la gente puede acceder a la realidad moral, no

tiene derecho a asumir que nuestras creencias morales de sentido común son fiables. Puesto que no se ha identificado ninguna ventaja de la explicación moral propuesta, aparte de su poder unificador, cualquier explicación igualmente sencilla que logre el mismo poder unificador debería ser igualmente aceptable. Así, aparentemente tendría el mismo sentido proponer que Malo = (*F* o *G* o *H*). O ¿por qué no inventar algún otro concepto no evaluativo, digamos "torfness", y decir que Torf = (*F* o *G* o *H*) -por qué no, en otras palabras, decir que (*F* o *G* o *H*) es idéntico a una sola propiedad, pero que esa propiedad no es evaluativa?

Los reduccionistas sintéticos admiten que el significado de un término evaluativo no es captado por ninguna expresión no evaluativa. Lo que ilustran las preguntas del último párrafo es que el significado distintivo y positivamente evaluativo de "bueno" no desempeña ningún papel en el tipo de explicaciones que nos ofrecería el reduccionista, ya que lo único que importa es que *F*, *G* y *H* se agrupen bajo un único concepto, un concepto negativamente evaluativo o un concepto no evaluativo servirían igual de bien. Esto contrasta con las explicaciones científicas legítimas. Cuando la hipótesis de que el agua = H_2O explica el resultado de la electrólisis del agua, el significado específico de "H_2O" ("un compuesto de dos átomos de hidrógeno y uno de oxígeno") desempeña un papel esencial. No se puede sustituir un concepto totalmente distinto, como "NaCl", en la teoría y obtener una explicación igual de buena: la hipótesis "agua = NaCl" no explica el resultado experimental.

4.5 El argumento de la disimilitud radical

Un argumento contra la reducción ética al que se ha prestado poca atención es el de que, a partir de nuestra comprensión de los conceptos

evaluativos, podemos ver simplemente la falsedad de las teorías reduccionistas. A primera vista, por ejemplo, lo incorrecto parece ser una propiedad completamente diferente de, por ejemplo, pesar 5 libras. En resumen:

1. Las propiedades del valor son radicalmente diferentes de las propiedades naturales.[46]
2. Si dos cosas son radicalmente diferentes, entonces una no es reducible a la otra.
3. Por tanto, las propiedades del valor no son reducibles a las propiedades naturales.

Si este argumento le parece poco convincente, probablemente sea porque acepta una de las dos respuestas siguientes.

En primer lugar, un reduccionista podría decir que no podemos ver la naturaleza de una cosa simplemente porque tenemos un concepto de ella. No podríamos ver, simplemente por nuestro concepto de agua, que el agua es idéntica al H_2O. Pero mi argumento no asume que podamos ver todo sobre la naturaleza de las cosas de las que tenemos conceptos; sólo asume que podemos ver algunas cosas sobre ellas, y que en el caso de las propiedades físicas y psicológicas, podemos ver que no son evaluativas. Para ilustrarlo, supongamos que un filósofo propone que el planeta Neptuno es la Novena Sinfonía de Beethoven. Creo que podemos ver que eso es falso, simplemente en virtud de nuestro concepto de Neptuno y nuestro concepto de sinfonías. Neptuno es una cosa totalmente diferente a la Novena Sinfonía de Beethoven. No hace falta ningún otro argumento. De hecho, si una persona no pudiera ver que Neptuno no es una sinfonía, diríamos que o bien no tiene idea de lo que es Neptuno, o bien no tiene idea de lo que es una sinfonía.

En segundo lugar, un reduccionista podría decir que el argumento, si

se aceptara, también "refutaría" la teoría de que el Agua =H_2O, ya que "agua" y "H_2O" parecen muy diferentes a primera vista. Del mismo modo, el argumento vería en refutar la teoría de que el sonido consiste en ondas de compresión en el aire, porque "sonido" y "ondas de compresión" parecen muy diferentes.

Pero no es así. A primera vista, el agua no parece ser H_2O; pero no es el caso que el agua parezca no ser H_2O. El concepto precientífico del agua la considera un líquido claro, inodoro, insípido, etc. Sobre la base de este concepto, no podemos discernir nada sobre cómo podría ser su microestructura. No tenemos la experiencia de ver simplemente que no hay partículas diminutas que la componen, etc. Tampoco tenemos la sensación de que el agua sea una categoría de cosa diferente del H_2O; por el contrario, vemos que son la misma categoría de cosa, es decir, sustancias físicas.

Tampoco tenemos la experiencia de ver que el sonido no consiste en ondas de compresión, o que son categorías de cosas diferentes. Podemos ver que son la misma categoría de cosas: eventos físicos. Sugiero que si tienes la intuición de que el sonido obviamente no puede ser una onda de compresión en el aire, entonces es porque por "sonido" en realidad te refieres a la experiencia del sonido, que realmente es un tipo de cosa diferente de las ondas de compresión en el aire.

El punto sobre la categoría amplia a la que pertenece una cosa es importante, porque en mi opinión, el hecho de que tengamos, al menos a grandes rasgos y en su mayor parte, intuiciones o creencias correctas sobre la naturaleza de *x* es una condición previa para que nos refiramos a *x*. Puedo hablar de Plotino sin saber mucho sobre él; lo único que sé es que fue un filósofo que vivió hace mucho tiempo. Pero no podría hablar de Plotino si tuviera creencias completamente inexactas (supuestas) sobre él; por ejemplo, si pensara que "Plotino" se refiere a un pastel,

entonces no estaría hablando de Plotino (la persona) en absoluto.[47] Del mismo modo, si "bueno" nos parece que se refiere a algo de una categoría fundamentalmente diferente de las propiedades naturales, entonces cuando decimos "bueno" no estamos hablando de una propiedad natural.

4.6 Explicación de las creencias morales

En las secciones 4.2 y 4.5 he demostrado que las propiedades evaluativas no pueden reducirse a propiedades naturales. En las secciones 4.3 y 4.4, he mostrado que el conocimiento moral no es empírico. La estructura de este último argumento era la siguiente:

1. Las afirmaciones morales se conocen empíricamente sólo si se conocen por observación o por razonamiento deductivo o científico a partir de las observaciones. (Premisa.)
2. Las afirmaciones morales no se conocen por observación. (§4.4.1)
3. Las afirmaciones morales no se conocen por deducción de premisas no morales. (§4.3)
4. Las afirmaciones morales no se conocen mediante el razonamiento científico a partir de premisas no morales. (§4.4)
5. Por lo tanto, las afirmaciones morales no se conocen empíricamente. (De 1, 2, 3, 4.)

Quiero hacer una última observación general sobre las perspectivas del reduccionismo en la ética. La razón por la que creemos que fenómenos como el calor, el agua y el sonido son reducibles es que tenemos argumentos científicos convincentes para identificar estos fenómenos con entidades teóricas específicas. Nadie duda de la existencia de las entidades teóricas relevantes, debido al poder sistemático de las teorías

científicas a las que pertenecen para explicar y predecir una serie de fenómenos observables. Sin esas teorías, nadie creería en el reduccionismo sobre el calor, el agua o el sonido.

En ética, no disponemos de teorías comparables. Los reduccionistas suelen limitarse a predecir que las afirmaciones morales resultarán reducibles.[48] Teorizan en abstracto sobre cómo sería posible justificar las creencias morales, pero sin presentar un argumento detallado para ninguna creencia moral específica. De hecho, los filósofos morales han trabajado durante dos milenios sin lograr un consenso sobre ninguna teoría moral general. En tal situación, la postura del reduccionista plantea dos cuestiones. La primera y más obvia es por qué deberíamos creer que es posible una reducción de los términos morales.

La segunda es cómo explica el reduccionista las creencias morales reales de la gente. Incluso si una de las propuestas discutidas en la sección 4.4 fuera viable, ninguna de ellas corresponde a cómo formamos realmente las creencias morales. Si "bueno" es un término teórico como "H_2O", ¿no deberíamos esperar que el concepto se desarrollara sólo al mismo tiempo que la teoría de la que forma parte, y que se dieran argumentos identificables antes de que se aceptara esa teoría? Cualquier químico puede identificar ejemplos incontrovertibles del tipo de observaciones que la teoría molecular debe explicar. ¿Por qué hay tanta gente incapaz de identificar ninguna de las observaciones que sus creencias morales deben explicar? Muchos profesionales de este campo niegan que existan tales observaciones, e incluso los reduccionistas éticos admitirán probablemente que el razonamiento empírico que subyace a las teorías morales es sutil y aún no se ha articulado con claridad. Sin embargo, las creencias morales abundan en todas las sociedades del mundo, entre jóvenes y mayores, educados e incultos. ¿Por qué? En lugar de declarar que casi todas las creencias morales

carecen todavía de justificación, ¿no deberíamos buscar algún medio mucho más sencillo de adquirir conocimientos morales?

Notas

1. G. E. Moore (1960, p. 40) y Michael Smith (1994, p. 17) caracterizan los fenómenos naturales como aquellos que pertenecen al ámbito de las ciencias naturales y/o de la psicología. Pero esto parece al revés: presumiblemente es la naturalidad de un fenómeno lo que hace que su estudio cuente como "ciencia natural", y no al revés. En otro lugar, Moore define los fenómenos "naturales" como aquellos que existen fuera del tiempo (PP. 40-1), pero no está claro por qué Moore pensaba que la bondad estaba fuera del tiempo mientras que las propiedades físicas estaban en el tiempo. En cualquier caso, no creo que ninguna de estas caracterizaciones capte la esencia de la disputa entre reduccionistas e intuicionistas en metaética.

2. Pretendo que esta caracterización del punto de vista reduccionista incluya la posibilidad de que sólo exista un relato muy complejo y disyuntivo -quizás incluso infinitamente complejo- de la naturaleza de la bondad. También pretendo incluir la opinión de que la bondad debería definirse en términos de alguna otra noción evaluativa (como la "razón"), siempre que se sostenga que el concepto o conceptos evaluativos fundamentales son reducibles a otros no evaluativos. Obsérvese que, por ejemplo, el relato de Scanlon sobre el valor no cuenta como una forma de reduccionismo ético en mi caracterización, ya que no considera que la noción normativa de "razones" sea reducible (1998, pp. 57-8, 95-100).

3. Autoproclamados naturalistas como Railton (19116, pp. 171-2), Michael Moore (1992, pp. 2498) y Sturgeon (1985, p. 51) han hecho de la posibilidad del conocimiento moral empírico un principio central de sus puntos de vista metaéticos.

4. Los argumentos de las secciones 4.2 y 4.4 siguientes están parafraseados de mi (2000a).

5. G. E. Moore (1960, pp. 17-19, 15-16) discute ambas formas de naturalismo, atribuyendo la primera a Jeremy Bentham (véase Bentham [1781]1996, p. 13).

6. Moore (1960, pp. 5-21).

7. Este punto procede de James Rachels (en conversación). Véase más arriba, sección

3.5, para una mayor exposición y crítica de las teorías del observador ideal.

8. Frankena (1973, pp. 99-100); Harman (1977, pp. 19-20).

9. El argumento es de Gettier (1963). Hasta donde yo sé, Weatherson (2003) es el único filósofo que sostiene que el análisis de la creencia justificada-verdadera puede ser aún correcto. Además, probablemente ningún filósofo ha sostenido que el argumento de Gettier fracasa debido al question-begging.

10. Frankena (1973, p. 99); Darwall, Gibbard y Railton (1992, p. 115).

11. El punto aquí es que, mientras que el análisis hedonista del "bien" puede parecer plausible cuando se plantea en abstracto, podemos identificar aplicaciones específicas de ese análisis a la luz de las cuales el análisis se vuelve inverosímil -de la misma manera que Gettier fue capaz de identificar una aplicación específica del análisis de la creencia justificada-verdadera del conocimiento a la luz de la cual ese análisis se volvió mucho menos plausible. Así, puede parecer verosímil que la bondad de la amistad consista en que conduzca al placer; pero no es verosímil que la bondad del placer consista en que conduzca al placer. Por supuesto, la analogía con Gettier no debe llevarse demasiado lejos: el argumento mooreano no implica, como el de Gettier, encontrar un caso en el que el analysans se aplique a algo mientras que el analysandum no.

12. Hay una cuestión complicada que no puedo abordar plenamente aquí, a saber, por qué los argumentos de la pregunta abierta no pueden utilizarse para refutar todas las definiciones. Por ejemplo, parece que "¿Es lo mismo un soltero que una soltera?" no es sinónimo de "¿Es lo mismo un soltero que una soltera?". ¿Por qué esto no demuestra que "soltero" no significa "hombre soltero"? Creo que la respuesta tiene que ver con el hecho de que algunas sentencias se refieren oblidagamente a expresiones, aunque no empleen los marcadores más obvios de mención de expresiones, como encerrar las expresiones pertinentes entre comillas. Creo que "¿Es lo mismo un soltero que un hombre soltero?" hace algo así como melltionillg (en lugar de o además de usar) las expresiones "soltero" y "hombre soltero", y eso bloquea la sustitución de "soltero" por "hombre soltero". Pero no creo que "¿Es bueno hacer lo que aumenta la cantidad de disfrute en el mundo?" mencione la expresión "bueno" o "lo que aumenta la cantidad de disfrute en el mundo". En cualquier caso, el argumento de la pregunta abierta debe formularse en términos de oraciones que no mencionan (sino que sólo utilizan) la expresión evaluativa pertinente y la frase de la definición naturalista propuesta.

13. Esta propuesta es de Rand (1964, p. 17).

14. Bume ([1739]1992, p. 469).

15. Asumo que la negación de cualquier afirmación evaluativa es una afirmación evaluativa, y que toda afirmación evaluativa es parte de algún sistema de valores. Aquí y en otros lugares, sólo considero las afirmaciones evaluativas y los sistemas de valores que son consistentes y no tautológicos. Aquí, el nihilismo es el punto de vista de que nada es bueno, malo, correcto o equivocado (esto difiere del uso en la sección 4.3.5).

16. Como ha señalado Stuart Rachels (comunicación personal), uno podría, en lugar de plantear una o más premisas evaluativas suprimidas, interpretar el argumento como si se basara en una regla evaluativa de inferencia, en la línea de "De la premisa de que *x* causa un gran sufrimiento, infiere que *x* es pro tanto malo". Tal vez sea una regla válida. Aun así, considero que seguiríamos enfrentándonos al problema de cómo estamos en condiciones de juzgar que esa regla es válida. Considero que la cuestión importante aquí es cómo conocemos las verdades morales, y parece que las reglas de inferencia del tipo en cuestión no nos ponen en posición de conocer las verdades morales a menos que estemos en algún sentido en posición de juzgarlas justificadamente como válidas. Sobre este último punto, véase Bonjour (2005, pp. 102, 116); Fumerton (de próxima aparición, capítulo 3).

17. Pigden (1991, p. 424) establece el principio de forma más general y precisa: "Ninguna expresión no lógica puede aparecer de forma no vacía en la conclusión de una inferencia válida a menos que aparezca en las premisas", donde se dice que una expresión aparece "vacía" si puede ser sustituida uniformemente por cualquier otra expresión del mismo tipo gramatical sin afectar a la validez de la inferencia. Pigden (1989, pp. 136-7) ofrece una prueba del principio así enunciado.

18. Pigden (1991, pp. 424-5) ha demostrado que la Ley de Hume es consistente con el reduccionismo.

19. Searle (1964, pp. 44-8).

20. Tal vez Searle cuestionaría esto; él representa la existencia de obligaciones como "hechos institucionales" creados en parte por reglas sociales que determinan lo que cuenta como ponerse bajo una obligación, al igual que las reglas del béisbol determinan, por ejemplo, lo que cuenta como un "home run" (1964, pp. 54-8). Creo que los argumentos contra el relativismo cultural expuestos en la sección 3.3 hacen que esta postura sea poco plausible. Zaibert (2003, p. 77) sostiene, sin embargo, que Searle no pretendía discutir las obligaciones morales en absoluto.

21. De Geach (1977) (con una ligera simplificación).

22. Véase Borowski (1980) y Hurka (1980) para otras críticas a Geach, aunque sólo encuentro convincente la de Hurka.

23. Prior (1976, pp. 90-1). Es en respuesta a estos ejemplos que Pigden introduce la calificación a la Ley de Hume mencionada en la nota 17.

24. Karmo (1988). A continuación he modificado ligeramente el argumento de Karmo para facilitar la exposición, y he añadido la Hipótesis de Determinación, que creo que Karmo necesita.

25. El argumento de Karmo se ha modificado ligeramente para facilitar la exposición y se ha añadido el supuesto de determinación, que, en mi opinión, Karmo necesita. La suposición de determinación podría cuestionarse mediante objeciones generales a la ley del término medio excluido, como las que se derivan de los problemas de vaguedad o de los nombres propios; sin embargo, tales objeciones no son relevantes en el contexto de la ética. Si se quiere, se puede estipular que 5 no debe contener expresiones vagas, nombres vacíos u otros fenómenos no propios de las oraciones éticas que puedan crear excepciones a la ley del medio excluido. Aunque algunos niegan en general que los enunciados evaluativos puedan ser verdaderos o falsos, esto es irrelevante aquí, ya que los reduccionistas éticos aceptan la capacidad de verdad de los enunciados evaluativos, y es a ellos a quienes se dirige el presente capítulo.

Mi declaración del principio de determinación incluye la frase "como máximo uno", en lugar de "exactamente uno", porque un mundo posible puede ser inconsistente con un sistema de valores. Por ejemplo, un sistema de valores que sostiene que todas las *F*'s son buenas y todas las *G* son malas es inconsistente con un mundo posible según el cual algunas *F* son *G*, aunque el sistema de valores y el mundo posible sean cada uno internamente consistente.

26. Maitzen (1998).

27. Maitzen (1998, p. 356).

28. El propio Maitzen (1998, p. 358) reconoce este hecho, aunque no saca la conclusión correcta de que el nihilismo es imposible.

29. Pero véase Horgan y Timmons (1992) para una nueva y, a mi juicio, persuasiva versión del Argumento de la Cuestión Abierta que se aplica al reduccionismo sintético.

30. Para las defensas de este punto de vista, véase Railton (1986), Sturgeon (1985),

Boyd (1988) y Michael Moore (1992).

31. Stuart Rachels (2003, p. 19) sugiere que uno puede observar que un caso de dolor en uno mismo es malo. Aunque esto es plausible, creo que el sentido de "observar" de Rachels es diferente al mío en esta sección: Me refiero a que los hechos evaluativos no pueden ser observados por los cinco sentidos, es decir, que no pueden ser vistos, oídos, saboreados, tocados u olidos. Sospecho que la "observación" de Rachels es lo que yo llamaría una intuición.

Al parecer, McGrath (2004) también ha rebatido aquí mi tesis; propone que a veces conocemos las verdades morales por percepción. Pero, de nuevo, no estoy seguro de que lo que tiene en mente sea la percepción en mi sentido, en lugar de la intuición. Véase especialmente su ejemplo ilustrativo en las páginas 224-5.

32. Harman (1977, pp. 4-5).

33. Harman (1977, p. 5): "[L]a pjercepción implica la formación de una creencia como resultado bastante directo de la observación de algo .... Si decimos que la observación se ha producido siempre que una opinión es un resultado directo de la percepción, debemos permitir que haya una observación moral ... "

34. Dos matizaciones: a) Puede haber varias formas distintivas del aspecto de las cosas *F*. b) También es suficiente que haya una forma de aspecto de las cosas *F*, y que el objeto observado no tenga ese aspecto. Ésta es sólo una condición necesaria para la observación. Para más información sobre la naturaleza de la observación, véase mi (2001, pp. 57-65, 93).

35. Una segunda objeción al relato implícito de Harman sobre la observación es que la observación no requiere creencia. Supongamos que veo una rata rosa, pero creo erróneamente que estoy alucinando. Entonces observo una rata rosa pero no creo que haya una rata rosa. Véase mi (2001, p. 93).

36. Moore (1992, p. 2517).

37. En lenguaje técnico, la cuestión es que el argumento de Michael Moore es falaz porque "*S* percibe que *x* es ___" es referencialmente opaco. Hay cierta ironía en que un reduccionista sintético cometa esta falacia, la misma falacia de la que los reduccionistas suelen acusar a G. E. Moore (véase Harman 1977, p. 19; Putnam 1981, pp. 206-8).

38. Sturgeon (1985, pp. 52, 54, 65).

39. Moore (1992, p. 2517).

40. Sturgeon (1985, p. 66).

41. La electrólisis del agua fue descubierta por William Nicholson en 1800. Lavoisier, Cavendish y otros ya habían argumentado en la década de 1780 que el agua es un compuesto de hidrógeno y oxígeno, basándose en experimentos en los que la combustión de hidrógeno en oxígeno produce agua pura. En el texto me he centrado en el experimento de la electrólisis por su efecto pintoresco.

42. Sturgeon (1985, pp. 51-2).

43. He modificado ligeramente el ejemplo de Sturgeon (1991, p. 29).

44. El ejemplo es de Putnam (1975, pp. 295-7).

45. Moore (1992, pp. 2530-1); compárese con Railton (1998, pp. 179-80).

46. Para los fines de este argumento, tal vez queramos restringir la atención a las propiedades evaluativas del paradigma, como la maldad, y a las propiedades naturales del paradigma, como pesar 130 libras. Se trata de propiedades cuyo estatus como evaluativas o naturales es claro e incontrovertible. En cambio, propiedades como ser placentero, ser sano y ser ilógico no son paradigmáticas; no está claro si son evaluativas, naturales, ambas o ninguna. La restricción a los paradigmas parece justa, ya que los naturalistas querrían sin duda reducir las propiedades morales paradigmáticas a las propiedades naturales paradigmáticas.

47. Compárese la objeción de Searle a la teoría causal de la referencia (1983, pp. 249-50) y la mía (2001, pp. 54, 60-2). Bealer (1999, sección 3) hace un punto relacionado, aproximadamente en el sentido de que la comprensión genuina de un concepto requiere disposiciones para tener intuiciones generalmente correctas que involucren el concepto. Mi opinión es que las disposiciones para tener intuiciones son cruciales para la comprensión de un concepto, más que las intuiciones o creencias reales. Sin embargo, las creencias pueden desempeñar un papel crucial en la referencia que uno hace a los particulares.

48. Sturgeon (1986, p. 73). Railton (1986, p. 190) es una excepción, ya que propone una reducción casi utilitaria de la rectitud.

Parte II:
Intuicionismo Ético

5

Conocimiento Moral

En los tres últimos capítulos hemos visto que las afirmaciones morales son afirmaciones sobre una clase de propiedades objetivas e irreductibles que no pueden conocerse a partir de la observación. ¿Cómo pueden conocerse estas afirmaciones, si es que pueden conocerse? ¿Es racional pensar que alguna de estas afirmaciones es verdadera? En el presente capítulo, explico cómo podemos conocer o estar justificados en la creencia de afirmaciones evaluativas sobre la base de la intuición ética.

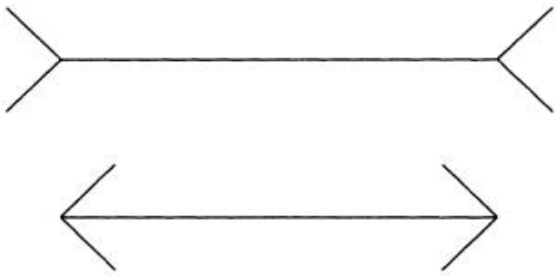
5.1 El principio del conservadurismo fenomenal

En igualdad de condiciones, es razonable suponer que las cosas son como parecen. Llamo a este principio "conservadurismo fenomenal" ("fenomenal" significa "perteneciente a las apariencias"). Ya he hablado de este principio en otro lugar, así que aquí seré relativamente breve.[1]

Hay un tipo de estado mental, que llamo "apariencia", que declaramos cuando decimos cosas como "me parece que *p*", "parece que

p" o "*p* es obvio", donde *p* es una proposición. Las apariencias tienen un contenido proposicional, es decir, lo que representan, pero no son creencias, como puede verse en la inteligibilidad de "El arco parece ser más alto que ancho, pero no creo que lo sea". Sin embargo, las apariencias nos llevan normalmente a formar creencias. La "apariencia" es una categoría amplia que incluye los estados mentales implicados en la percepción, la memoria, la introspección y la intelección. Así, podemos decir: "Esta línea parece más larga que aquella", "Me parece recordar haber leído algo sobre eso", "Me parece que me duele la cabeza" y "Parece que dos puntos cualesquiera pueden estar unidos por una sola línea recta".[2] Todas esas afirmaciones tienen sentido, utilizando el mismo sentido de "parece".

Las apariencias pueden ser engañosas, y las apariencias pueden entrar en conflicto entre sí, como en la ilusión de Müller-Lyer:



Al principio parece que la línea superior es más larga que la inferior. Pero si sacas una regla y las mides, verás que tienen la misma longitud. La línea superior parecerá, al sostener una regla junto a ella, que mide 5 centímetros, y la línea inferior parecerá igualmente que mide 5 centímetros. Por lo tanto, parece que las dos líneas tienen la misma longitud. Como ilustra este ejemplo, una apariencia inicial puede ser anulada por otras apariencias (esto no significa que la apariencia inicial desaparezca, sino sólo que no la creemos), y sólo por otras apariencias. Algunas apariencias son más fuertes que otras -como decimos, algunas

cosas son "más obvias" que otras- y esto determina a qué nos aferramos y qué rechazamos en caso de conflicto. Es probable que te parezca más claro que el resultado de medir las líneas es exacto que el resultado de verlas a simple vista, por lo que crees en el resultado de la medición (esto puede tener que ver con las creencias de fondo que tienes sobre la fiabilidad de los diferentes procedimientos, que a su vez se basan en la forma en que te parecen otras cosas). Las cosas pueden complicarse cuando intervienen muchas creencias y/o apariencias diferentes, pero el principio básico es que nos inclinamos más a aceptar lo que nos parece más verdadero.

Las apariencias pueden ser intelectuales, a diferencia de las sensoriales, mnenlónicas o introspectivas. Nos parece que el camino más corto entre dos puntos cualesquiera debe ser una línea recta; que el tiempo es unidimensional y totalmente ordenado (para dos momentos cualesquiera, uno es anterior al otro); y que ningún objeto puede ser completamente rojo y completamente azul al mismo tiempo. Acepto estas cosas por motivos intelectuales. No estoy mirando todos los pares de puntos posibles y todos los caminos posibles que conectan cada par y viendo, con mis ojos, que el camino recto es el más corto en cada caso. En cambio, estoy "viendo" intelectualmente que no es cierto, es decir, que cuando pienso en ello, se vuelve obvio.

Los juicios lógicos se basan en apariencias intelectuales. Pensamos que la siguiente inferencia es lógicamente válida (las premisas implican la conclusión, independientemente de que las premisas sean verdaderas):

> Sócrates es un hombre.
>
> Todos los hombres son desconsiderados.
>
> Por lo tanto, Sócrates es desconsiderado.

Pero la siguiente no es válida:

Sócrates es desconsiderado.

Todos los hombres son desconsiderados.

Por lo tanto, Sócrates es un ornitorrinco.

Nosotros "vemos" esto, no con nuestros ojos, sino con nuestro intelecto o razón.

Todos los juicios se basan en cómo le parecen las cosas al sujeto que las juzga: una persona racional cree sólo lo que le parece verdadero, aunque no tiene por qué creer todo lo que le parece verdadero.[3] La función de los argumentos es cambiar la forma en que las cosas le parecen a la audiencia, presentando otras proposiciones (premisas) que parecen verdaderas y que parecen apoyar algo (la conclusión) que inicialmente puede no haber parecido verdadero a la audiencia. Un argumento sólo tiene fuerza en la medida en que sus premisas parecen verdaderas y parecen apoyar su conclusión. La investigación intelectual presupone el Conservadurismo Fenomenal, en el sentido de que dicha investigación procede asumiendo que las cosas son como parecen, hasta que surgen pruebas (a su vez extraídas de las apariencias) que lo ponen en duda. Incluso los argumentos de un escéptico filosófico que dice que no está justificado creer en nada, se basan en las propias creencias del escéptico, que se basan en lo que al escéptico le parece verdadero.

Esto indica en pocas palabras por qué considero que cualquier negación del conservadurismo fenomenal es contraproducente. Sea como fuere, ya hemos establecido en el capítulo 1 que el escepticismo filosófico general está fuera de la mesa en la presente discusión. Dado que todo juicio y razonamiento presupone el conservadurismo fenomenal, un rechazo del conservadurismo fenomenal equivale a un escepticismo filosófico general. Por lo tanto, asumimos que el Conservadurismo Fenomenal es correcto.

5.2 Intuiciones éticas

El razonamiento a veces cambia la forma en que nos parecen las cosas. Pero también hay un modo en que las cosas nos parecen antes del razonamiento; de lo contrario, el razonamiento no podría iniciarse. La forma en que las cosas se ven antes del razonamiento puede llamarse "apariencia inicial". Una apariencia inicial, intelectual, es una "intuición". Es decir, una intuición de que *p* es un estado de su apariencia que no depende de la inferencia de otras creencias y que resulta de pensar en *p*, a diferencia de percibir, recordar o introspeccionar.[4] Una intuición ética es una intuición cuyo contenido es una proposición evaluativa.

Muchos filósofos se quejan de que no saben qué es una intuición o de que el término "intuición" es esencialmente vacío y no da cuenta de cómo se puede saber algo.[5] Supongo que estas críticas acaban de ser contestadas.

Algunos se preguntan si las intuiciones existen.[6] Hemos visto algunos ejemplos de intuiciones en la sección anterior. He aquí algunos ejemplos en ética:

> Disfrutar es mejor que sufrir.
>
> Si *A* es mejor que *B*, y *B* es mejor que *C*, entonces *A* es mejor que *C*.
>
> Es injusto castigar a una persona por un delito que no ha cometido.
>
> El valor, la benevolencia y la honestidad son virtudes.
>
> Si una persona tiene derecho a hacer algo, nadie tiene derecho a impedirle por la fuerza que lo haga.

Antes de considerar los argumentos a favor o en contra, cada una de

estas proposiciones parece verdadera. En cada caso, la apariencia es intelectual; no se percibe que estas cosas sean así con los ojos, los oídos, etc. Y son evaluativas.[7] Así que los estados mentales relevantes son intuiciones éticas.

He aquí algunos ejemplos de afirmaciones éticas que, entiendo, no son intuitivas, ni siquiera para quienes las creen:

> Estados Unidos no debería haber entrado en guerra en Irak en 2003.
>
> Deberíamos privatizar la Seguridad Social.
>
> El aborto está mal.

Aunque estas proposiciones parecen ciertas para algunos, las apariencias relevantes no cuentan como "intuiciones" porque dependen de otras creencias. Por ejemplo, la sensación de que Estados Unidos no debería haber invadido Irak depende de creencias como que la guerra causó previsiblemente miles de muertos, que eso es malo, que Irak no tenía armas de destrucción masiva, etc. Esto no quiere decir que la intuición tenga un papel en la conclusión de que Estados Unidos no debería haber ido a la guerra. Es la intuición la que nos dice que matar gente es, a primera vista, incorrecto. La intuición también interviene en la ponderación de valores contrapuestos: por ejemplo, podemos tener intuiciones sobre si es correcto matar a muchas personas para deponer a un tirano, si los hechos del caso son los que creemos.

Algunos piensan que las intuiciones son sólo creencias y que, por tanto, la "intuición" no nombra una forma de saber nada,[8] pues no queremos decir que por el mero hecho de creer algo, lo sé. Una preocupación más sofisticada es que lo que pensamos como intuiciones puede ser producto de creencias ya existentes, quizás a través de inferencias subconscientes. Quizás "disfrutar es mejor que sufrir" sólo me parece cierto porque ya lo creo, o porque creo cosas de las que se

deduce. Hay dos respuestas a estas preocupaciones. En primer lugar, la opinión de que las intuiciones son o están causadas por las creencias no explica el origen de nuestras creencias morales. Sin duda, algunas creencias morales se explican por inferencia de otras creencias morales. Pero como ninguna creencia moral puede derivarse de premisas totalmente no morales, debemos comenzar con algunas creencias morales que no se infieren de ninguna otra creencia. ¿De dónde proceden estas creencias morales de partida? ¿Las adoptamos de forma totalmente arbitraria? No; ésta no es la fenomenología de la creencia moral. Adoptamos creencias morales fundamentales porque nos parecen correctas; no las seleccionamos al azar.

En segundo lugar, las intuiciones morales no están causadas en general por las creencias morales previas, ya que las intuiciones morales a menudo entran en conflicto con nuestras teorías morales previas, o simplemente no se explican por ellas. He aquí dos famosos ejemplos hipotéticos de la literatura ética:

> *Ejemplo* 1: Un médico de un hospital tiene cinco pacientes que necesitan trasplantes de órganos; de lo contrario, morirán. Todos ellos necesitan órganos diferentes. También tiene un paciente sano, que acude a una revisión rutinaria y que resulta ser compatible con los cinco. ¿Debe el médico matar al paciente sano y distribuir sus órganos entre los otros cinco?

> *Ejemplo* 2: Un carro desbocado se dirige a una bifurcación de la vía. Si toma la bifurcación de la izquierda, chocará con cinco personas y las matará; si toma la bifurcación de la derecha, chocará con una persona y la matará. Ninguna de las personas puede apartarse a tiempo. Hay un interruptor que determina qué bifurcación toma el carro. Actualmente está configurado para enviar el carro a la izquierda. Puedes

> activar el interruptor, enviando el carro a la derecha en su lugar. ¿Hay que darle la vuelta al interruptor?[9]

La mayoría de la gente responde intuitivamente "no" al ejemplo 1 y "sí" al ejemplo 2. Algunos filósofos sostienen que la acción moralmente correcta es siempre la que tiene las mejores consecuencias generales. Su punto de vista implica que la respuesta al ejemplo 1 es "sí". Pero incluso estos filósofos, cuando se enfrentan al ejemplo, admiten que su respuesta es contraintuitiva, que se ve mal matar al paciente sano y extraer sus órganos.[10] Las intuiciones de uno no se ajustan simplemente a lo que uno cree sobre la moralidad. En este sentido, la mayoría de las personas tienen dificultades para explicar por qué se sienten inclinadas a responder de una manera en el ejemplo 1, y de la manera opuesta en el ejemplo 2; ambos casos introducen la posibilidad de sacrificar a una persona para salvar a cinco. Los filósofos han propuesto varias explicaciones al respecto (que siguen siendo controvertidas). La cuestión es que ninguna teoría moral mantenida antes de considerar casos como los anteriores puede ofrecernos una explicación de por qué el sacrificio debería considerarse inaceptable en el ejemplo 1 pero aceptable en el ejemplo 2.

El hecho de que la intuición sea a menudo independiente de la creencia es importante, ya que permite a la intuición proporcionar el tipo de restricción necesaria para decidir entre teorías morales opuestas. Si la intuición se limitara a seguir las creencias morales, no podría ayudarnos a decidir qué creencias morales son correctas. Pero este punto es compatible con que la intuición muestre cierto grado de receptividad a nuestras creencias, y no quiero afirmar que las intuiciones de una persona permanezcan en general totalmente ajenas a las teorías que adopta. Compárese con la observación de que las percepciones sensoriales son en gran medida, pero no totalmente, independientes de

nuestras creencias de fondo: por ejemplo, aunque yo crea que Pie Grande no existe, si Pie Grande se me acerca, lo veré.

Entre las proposiciones morales intuitivas, algunas son más intuitivas que otras. Compare los dos ejemplos anteriores con los siguientes:

> *Ejemplo* 3: Como en el ejemplo 2, salvo que no hay nadie en la bifurcación de la derecha; si el carro baja por la bifurcación de la derecha, chocará con un montón de arena que lo detendrá con seguridad. ¿Debes accionar el interruptor?

Todo el mundo responde "sí" a esta pregunta, incluso los que respondieron "no" al ejemplo 2. Nuestras intuiciones sobre el ejemplo 3 son más claras y seguras que las de los ejemplos 1 y 2. Esto hace que la creencia de que hay que pulsar el interruptor en el ejemplo 3 tenga un mayor nivel de justificación que las creencias correspondientes sobre los ejemplos 1 y 2.

Al oír estos ejemplos, algunas personas intentan negar las intuiciones que he señalado planteando problemas como: "En el ejemplo 1, ¿y si la paciente sana es la futura madre de Josef Stalin?" "En el ejemplo 3, ¿y si las cinco personas de la horquilla izquierda son suicidas que fueron allí para ser atropellados y se van a buscar otra forma de suicidarse?". Y al considerar la intuición de que el disfrute es mejor que el sufrimiento: "¿Y si el disfrute es un placer sádico o pervertido?". La respuesta a todas estas preguntas es la misma: estipulo que esas cosas no son así. En todos mis ejemplos, todas las condiciones deben suponerse normales a menos que se especifique lo contrario; del mismo modo, la mayoría de los principios morales llevan implícita la cláusula "en condiciones normales". El propósito de considerar estos ejemplos no es iniciar un ejercicio legalista de búsqueda de lagunas en un enunciado y de formas de llenar dichas lagunas. Nuestro objetivo actual es simplemente mostrar la existencia y la naturaleza de las intuiciones éticas.

No todas las intuiciones son iguales: algunas son más creíbles que otras. Como sugieren las observaciones anteriores, una razón para ello es que algunas intuiciones son simplemente más fuertes, o parecen más claramente verdaderas, que otras. Otra razón es que algunas intuiciones son más compartidas que otras; en igualdad de condiciones, una intuición con la que muchos no están de acuerdo tiene más probabilidades de ser un error que una intuición que casi todos comparten. Otra razón es que algunas intuiciones tienen un contenido más sencillo que otras y, por tanto, son menos propensas al error. Y hay varias razones por las que algunos tipos de intuiciones pueden ser más propensos a los sesgos que otros. Estos hechos apuntan a la conclusión de que las intuiciones no deben aceptarse de forma acrítica, y que las intuiciones conflictivas deben sopesarse entre sí teniendo en cuenta nuestros mejores juicios en cuanto a sus niveles relativos de fiabilidad. Volveré sobre este punto en el siguiente capítulo, cuando abordemos la cuestión de la resolución de los desacuerdos éticos (secciones 6.4 y 6.6).

5.3 Malentendidos del intuicionismo

Las proposiciones intuitivas que hemos estado discutiendo están justificadas *prima facie*. Es decir, está justificado que las creamos a menos que surja una evidencia contraria lo suficientemente fuerte como para derrotar la presunción inicial a su favor. Estas pruebas contrarias consistirían en pruebas directamente contrarias a la proposición que intuitivamente parecía verdadera, o en pruebas de que nuestra intuición inicial no era fiable.

Ahora podemos ver que al menos una objeción al intuicionismo se basa en una interpretación errónea de la doctrina. Karl Popper escribe:

El "intuicionismo" es el nombre de una escuela filosófica que enseña

> *que tenemos alguna facultad o capacidad de intuición intelectual que nos permite "ver" la verdad; de modo que lo que hemos visto como verdadero debe serlo. Se trata, pues, de una teoría de una fuente de conocimiento autorizada.*

A continuación, critica el intuicionismo porque las intuiciones pueden ser erróneas y debemos estar abiertos a revisar nuestras opiniones éticas.[11] Es de suponer que piensa que los intuicionistas niegan esas cosas, pero pocos intuicionistas, si es que hay alguno, lo han hecho, y no hay ninguna razón para que lo hagan.[12] El mismo malentendido puede estar detrás de la acusación ocasional de que el intuicionismo es "dogmático".[13] No quiero descartar (como hace Popper) la posibilidad de que algunas intuiciones sean infalibles; simplemente niego que deban ser infalibles.

Tara Smith malinterpreta el intuicionismo como la opinión de que todas las verdades morales son "evidentes por sí mismas". De hecho, los intuicionistas sostienen como mucho que algunas verdades morales son evidentes por sí mismas,[14] y mi propia forma de intuicionismo sostiene únicamente que algunas creencias morales se justifican *prima facie* por intuiciones. Así pues, no se genera ningún problema para el intuicionismo por citar ejemplos de principios morales que se basan en el razonamiento, ni por citar principios morales que no son 100% ciertos. El intuicionismo tampoco afirma la "irrelevancia del argumento" en general.[15] Una vez que tenemos un fondo de creencias morales justificadas *prima facie* del que partir, hay un gran margen para que el razonamiento moral amplíe, refine e incluso revise nuestras creencias morales, exactamente de la manera que muestra la literatura contemporánea sobre ética filosófica.

Hay que reconocer que los críticos del intuicionismo no se han quedado sin excusa en los malentendidos mencionados.[16] H. A. Prichard,

una de las principales figuras del intuicionismo del siglo XX, al menos los invitó, y quizás en su caso ni siquiera eran malentendidos:

> *Esta realización de la autoevidencia [de nuestras obligaciones] es un conocimiento positivo, y en la medida en que el término Filosofía Moral se limita a este conocimiento y al conocimiento de la inmediatez paralela de la aprehensión de la bondad de las diversas virtudes y de las buenas disposiciones en general, existe tal cosa como Filosofía Moral.*[17]

El uso que hace del término "autoevidencia" favorece la lectura de Popper (aunque en realidad dice que lo único que quiere decir con "autoevidencia" es "no-derivada"), y el resto del pasaje favorece la lectura de Tara Smith. Pero -dejando de lado la cuestión interpretativa- un filósofo que discute una teoría debería abordar la versión más fuerte de la teoría, no la más débil. Concediendo la justificación, sobre la base de la intuición, de los principios morales de sentido común, no hay ninguna motivación derivada de ningún supuesto central del intuicionismo para negar que la filosofía moral pueda construir más argumentos, llegando a verdades morales no inmediatamente evidentes. La analogía que Prichard establece con las matemáticas debería sugerirnos, en todo caso, que los elementos derivados del conocimiento ético podrían superar con creces a los intuitivos.

Algunos pueden pensar que la fundamentación del intuicionismo requiere una doctrina de la infalibilidad: es decir, la idea de que podemos partir de algunos principios morales, sin tener que justificarlos con argumentos, implica que esos principios morales deben ser infalibles, incorregibles, o algo parecido. Nunca he conseguido que nadie me diga por qué esto sería así.[18] ¿Por qué no podemos mantener nuestros puntos de partida abiertos a la revisión en caso de que surjan tensiones con otras creencias justificadas? Supongamos que me parece ver un vaso de agua

sobre la mesa. Eso es suficiente para que esté justificada mi creencia de que hay un vaso de agua, en ausencia de cualquier evidencia contraria. Sin embargo, puedo seguir creyendo que hay un vaso de agua, pero si al cogerlo me doy cuenta de que mi mano lo atraviesa, y si una docena de personas en la sala dicen que no hay ningún vaso, puedo decidir que no había ningún vaso. Como ilustra este ejemplo, normalmente consideramos que las creencias perceptivas están justificadas a primera vista, tal y como dicta el principio del conservadurismo fenoménico. No hay ningún obstáculo evidente para considerar que las creencias intuitivas están justificadas de forma similar.

5.4 Objeciones epistemológicas comunes

Objeción 1:

Necesitamos razones para creer que nuestras intuiciones éticas, o la facultad de intuir en general, son fiables. De lo contrario, las intuiciones no pueden justificar nuestras creencias morales.[19]

Respuesta:

¿Qué sucede si aplicamos el principio de manera general: "Necesitamos razones positivas para confiar en las apariencias"? Entonces necesitamos razones positivas para confiar en la percepción sensorial, en la memoria, en la introspección, incluso en la propia razón. El resultado es el escepticismo global. No se puede aceptar nada hasta que no demos una razón positiva para confiar en ese tipo de creencia. Pero no podemos dar tal razón sin basarnos en la percepción de los sentidos, la memoria, la introspección, la razón -o, en general, en alguna fuente-. Por lo tanto, nunca podremos confiar en nada.[20] Por supuesto, esto significa que tampoco podríamos confiar en el razonamiento de este

párrafo.

Hemos estipulado que el escepticismo filosófico general no es nuestra preocupación. No nos interesa aquí discutir la opinión de que nadie puede conocer las verdades morales porque nadie puede conocer nada en absoluto. Se podría intentar evitar la amenaza escéptica recurriendo a una teoría de la coherencia de la justificación, según la cual las creencias se justifican por sus relaciones de apoyo mutuo entre sí, en lugar de construirse a partir de fundamentos justificados independientemente. En mi opinión, hay objeciones convincentes a tal teoría, pero no puedo discutirlas aquí.[21] Para los propósitos actuales, baste decir que si una teoría de este tipo puede tener éxito en la justificación de nuestras otras creencias, no hay ninguna razón aparente por la que no pueda también reivindicar las creencias morales. Las creencias morales pueden apoyarse mutuamente tan bien como cualquier otro tipo de creencia. Uno podría preocuparse por cómo la intuición moral se integraría en dicha teoría, pero también podría preocuparse por cómo la percepción se integraría en la teoría de la coherencia. Si el coherentista puede acomodar de alguna manera el papel de la percepción en la justificación de nuestras creencias empíricas, no está claro por qué no podría acomodar la intuición de manera similar.

Pero no creo que los defensores de esta primera objeción pretendan respaldar ni el coherentismo ni el escepticismo en epistemología. Más bien creen que la intuición es de algún modo especial, de manera que la someten a una exigencia general de fundamentos justificativos, una exigencia de la que están exentos la percepción, la memoria, la introspección y el razonamiento. En vista del principio del conservadurismo fenomenal, no está claro por qué debería ser así; las intuiciones son sólo otro tipo de apariencia, junto con las experiencias perceptivas, las experiencias de la memoria, etc. Además, en la sección

5.1 vimos ejemplos de intuiciones no morales que, supongo, casi todo el mundo aceptaría. Si uno acepta esas intuiciones, parecería arbitrario no aceptar también las intuiciones éticas, al menos *prima facie*.

Objeción 2:

El problema de las intuiciones es que nunca podemos comprobar si una intuición ética es correcta, sin basarnos en la intuición.[22] En cambio, las creencias empíricas pueden comprobarse a menudo por otros medios. Si dudo de que la mesa que veo sea real, puedo comprobarlo intentando tocarla, preguntando a otras personas si también la ven, o intentando poner un vaso sobre ella.

Respuesta:

Hay tres respuestas a esta objeción. En primer lugar, la objeción suena sospechosamente como la Objeción 1. Si consideramos que las creencias están justificadas *prima facie* sobre la base de las apariencias, entonces no está claro por qué debería pensarse que las creencias intuitivas requieren comprobación, en ausencia de cualquier motivo positivo para dudar de ellas. Si, por otro lado, rechazamos esta concepción de la justificación prima facie, entonces no está claro cómo se supone que hay que comprobar nada. Si la creencia A no tiene justificación *prima facie* y la creencia B tampoco la tiene, entonces no se puede "comprobar" o "verificar" legítimamente la verdad de A apelando a B. A menos que se nos permita dar algo por sentado, nada puede considerarse como verificación de nada.

En segundo lugar, es dudoso que todo nuestro conocimiento no moral pueda ser comprobado en el sentido requerido por la objeción. Creo que tengo estados mentales -creencias, deseos, sentimientos, etc.- porque (parece) que tengo conciencia introspectiva de ellos. No estoy seguro de

cómo podría comprobar la fiabilidad de la introspección por medios no introspectivos, y no creo que lo haya hecho nunca. Sin embargo, es bastante seguro que tengo estados mentales. Del mismo modo, no está claro cómo podría comprobar la fiabilidad general de la memoria, sin basarme en la memoria; la fiabilidad del razonamiento inductivo, sin basarme en la inducción; o la fiabilidad de la razón en general, sin basarme en la razón.[23] Incluso los ejemplos citados en el enunciado de la objeción podrían no contar como comprobación de una creencia empírica por otros medios: si la creencia de que hay una mesa aquí se clasifica como basada en la "percepción sensorial", entonces todos los medios sugeridos para verificar la creencia se basan en la misma fuente. Por lo tanto, esta objeción corre el peligro de convertirse en un escepticismo filosófico general.

En tercer lugar, si se adopta un punto de vista liberal sobre lo que cuenta como comprobación de una creencia -como debe hacerse para permitir que la mayoría de las creencias no morales sean "comprobadas"-, parece que las intuiciones pueden ser comprobadas. Puedo comprobar mi creencia de que el asesinato está mal preguntando a otras personas si el asesinato también les parece mal. Si es legítimo, como seguramente lo es, comprobar una creencia perceptiva de este modo, ¿por qué no debería ser una comprobación igualmente válida de una creencia intuitiva? También se puede comprobar una creencia intuitiva viendo si es coherente con otras creencias intuitivas, al igual que se puede comprobar una creencia perceptiva viendo si es coherente con otras creencias perceptivas. Así, supongamos que alguien tiene la intuición de que el aborto es un error. Puede comprobarlo (a) viendo si su intuición es coherente con las intuiciones de otros, y (b) viendo si su intuición sobre el aborto es coherente con su intuición sobre, por ejemplo, el caso del violinista de Thomson.[24] Este tipo de pruebas no son triviales: muchas intuiciones no las superan, aunque muchas otras sí. No

es que el intuicionista remita inmediatamente todas las cuestiones morales a la intuición, sin posibilidad de discusión o razonamiento posterior.

Objeción 3:

Si permitimos que las creencias morales se apoyen en meras apelaciones a la intuición, entonces cualquiera puede afirmar que cualquier creencia moral está justificada. Si Thelma puede estar justificada no inferencialmente en la creencia de que comer carne es malo, entonces Luisa también puede estar justificada no inferencialmente en la creencia de que comer carne no es malo, incluso si ninguno de los dos puede inferir su creencia a partir de ninguna razón".[25]

Respuesta:

Cuando uno percibe un objeto físico, a primera vista está justificado creer algunas cosas sobre el objeto, cosas que se pueden discernir perceptualmente. De ello no se deduce que cualquier afirmación arbitraria sobre el mundo físico esté justificada. Del mismo modo, sostengo que cuando uno tiene una intuición ética, está *prima facie* justificado para creer en la proposición evaluativa relevante; de esto no se deduce que cualquier proposición evaluativa elegida arbitrariamente esté justificada.

Quizá la cuestión es que Luisa estaría justificada al pensar que comer carne no está mal, si tuviera la intuición ética correspondiente. Por supuesto, esto se deduce de mi teoría. También es cierto que si alguien mirara al cielo y tuviera una experiencia visual de enrojecimiento, a primera vista estaría justificado que creyera que el cielo es rojo. ¿Cuál es el problema?

Quizás la objeción se basa en la suposición de que muchas personas tienen de hecho la intuición de que comer carne no es malo. Esto sería un problema para alguien que quiere mantener que comer carne es malo, igual que sería un problema para alguien que cree que el cielo es azul si muchas personas miran hacia arriba y ven colores diferentes. Si esta es la objeción, entonces cae bajo el epígrafe del argumento del desacuerdo, que se discutirá en el capítulo 6.

Algo que no supone un problema para el intuicionista es la posibilidad de que haya personas que afirmen indiscriminadamente que tienen intuiciones, quizá porque no les apetece exponer las verdaderas razones de sus creencias. No tenemos ninguna técnica general para obligar a la gente a ser sincera y cuidadosa. Esto es lamentable. Pero no influye en la realidad de la intuición ni en su validez como fuente de conocimiento. Del mismo modo, los testigos presenciales pueden exagerar, hacer juicios precipitados y mentir. Nadie cree que esto refute la validez de la percepción sensorial como medio de conocimiento. Tampoco encargamos al filósofo de la percepción la tarea de impedir que la gente haga esas cosas. Por lo tanto, no es la tarea del intuicionista ético producir una técnica para obligar a todos a ser circunspectos y honestos en sus afirmaciones de valor.

Objeción 4:

John Mackie llama a la intuición ética "extraña" y "completamente diferente de nuestras formas ordinarias de conocer todo lo demás". Ninguno de nuestros relatos ordinarios sobre la percepción sensorial o la introspección, o la formulación y confirmación de hipótesis explicativas, o la inferencia, o la construcción lógica, o el análisis conceptual, o cualquier combinación de éstos" puede explicar el conocimiento ético.[26]

Respuesta:

Dada la realidad de la intuición en general, la intuición ética no es muy diferente de otros tipos de intuición. La única diferencia entre las intuiciones éticas y las no éticas está en lo que son, y eso no puede ser tomado como base para la rareza que Mackie ve, a menos que debamos rechazar el conocimiento ético simplemente por ser ético.

Sin duda, Mackie diría que es la intuición en general la que es extraña y totalmente diferente de otros medios de conocimiento. Está notablemente ausente de su lista de nuestras formas "ordinarias" de conocer las cosas. Pero no es un argumento contra el conocimiento intuitivo decir que no se puede explicar por ninguno de los medios no intuitivos que tenemos para conocer las cosas.[27] También se podría argumentar que la percepción es rara, ya que el conocimiento perceptivo no puede ser explicado por la introspección, la intuición, el análisis conceptual o el razonamiento. El hecho es que Mackie no ha identificado ninguna característica específica de la intuición que la haga problemática. Uno sospecha que su referencia a la "rareza" del conocimiento moral carece de significado cognitivo, sirviendo más bien para expresar su propia aversión a tales cosas que para describir cualquier característica objetiva del mismo.[28]

Detrás de la aversión de Mackie por la intuición se esconde sin duda algo del fuerte sentimiento empirista de la filosofía del siglo XX. El empirismo -a grandes rasgos, la idea de que todo el conocimiento "informativo", o el conocimiento del mundo independiente de la mente y del lenguaje, debe derivar de la percepción de los sentidos- ha estado de moda durante el último siglo, aunque menos, creo, en la última década. No puedo hacer justicia a este tema aquí; sin embargo, informaré brevemente de cómo me parecen las cosas. En primer lugar, es tan fácil enumerar lo que a primera vista parecen ser contraejemplos de la tesis

del empirismo, y al mismo tiempo es tan difícil encontrar argumentos a favor de la tesis, que la motivación subyacente de la doctrina sólo puede asumirse como un prejuicio. En segundo lugar, creo que en los últimos años, si no antes, se ha demostrado que la doctrina es insostenible.[29] Aquí daré dos de los contraejemplos más conocidos del empirismo.

Primer ejemplo: Nada puede ser a la vez totalmente rojo y totalmente verde.[30] ¿Cómo puedo saber eso? Obsérvese que la cuestión no es cómo he llegado a los conceptos "rojo" y "verde", ni cómo he llegado a entender esta proposición. La cuestión es por qué, habiéndola entendido, estoy justificado para afirmarla, en lugar de negarla o retener el juicio. Parece estar justificada intuitivamente, es decir, simplemente porque parece obvia al reflexionar. ¿De qué otra forma podría estar justificada?

Un empirista ingenuo podría apelar a mis experiencias con objetos de color: He visto muchos objetos de color, y ninguno de ellos ha sido a la vez rojo y verde. Una cosa que hace que esto sea inverosímil como explicación de cómo sé que nada puede ser a la vez rojo y verde es la necesidad del juicio. Contrasta los dos siguientes estados:

> Nada es a la vez verde y rojo.
>
> Nada es a la vez verde y un millón de millas de largo.

Nunca hemos observado un ejemplo contrario a ninguna de las dos afirmaciones, por lo que parece que la segunda está al menos tan bien respaldada por la observación como la primera. La segunda afirmación es probablemente cierta, ya que nunca hemos observado un objeto verde que mida un millón de kilómetros, aunque no parece haber ninguna razón para que no pueda existir. Tenemos una idea clara de lo que sería observar una cosa así, y no sería insensato buscarla. Pero la primera afirmación es diferente: podemos ver que simplemente no podría haber un objeto verde que sea rojo, y parece que, independientemente de cómo

haya sido nuestra experiencia, no habríamos dicho que existiera tal objeto; en consecuencia, no tendría sentido ni siquiera buscarlo. Estos puntos son difíciles de cuadrar con el argumento de que ambas afirmaciones se justifican de la misma manera, por la mera falta de observación de un contraejemplo. Además, supongamos que todas o la mayoría de las observaciones de objetos de color han sido alucinatorias (quizás, como Neo, te enteras de que vives en Matrix). Según el presente relato empirista, tendrías que suspender el juicio sobre si, en el mundo real, los objetos rojos son también verdes. Esto parece absurdo.

Por esta razón, muchos de los que simpatizan con el empirismo se inclinan por afirmar que "Nada puede ser rojo y verde a la vez" se hace cierto en virtud de las definiciones de "rojo" y "verde". A menudo se piensa que ésta es una forma aceptable de evitar la dependencia de la intuición. Pero no basta con hacer este tipo de afirmación; para cumplirla, el empirista debe presentar las definiciones de "rojo" y "verde" junto con la derivación real, a partir de esas definiciones, de la afirmación "Nada puede ser tanto rojo como verde". Nadie lo ha hecho; de hecho, el proyecto parece estilizado en la primera etapa por la ausencia de cualquier definición analítica de "rojo" o "verde". Es aquí donde algunos se ven obligados a apelar a los conocimientos científicos sobre la naturaleza subyacente de los colores para construir definiciones (diciendo, por ejemplo, "el rojo es la disposición a retener tales y tales longitudes de onda de la luz"). Pero este enfoque lleva a la absurda consecuencia de que, digamos, hace 300 años, la gente no estaba en condiciones de saber si era posible que un objeto rojo fuera verde -de hecho, ni siquiera entendían el significado de esas palabras-, ya que no conocían la teoría científica de los colores.

Segundo ejemplo: Sé que "Sócrates es un hombre" y "Todos los hombres son machistas" implican conjuntamente que "Sócrates es

machista". ¿Cómo lo sé? Se podría decir que lo sé porque conozco una regla general según la cual todas las inferencias de la forma "x es una A; todas las A son B; por tanto, x es B" son lógicamente válidas, pero, en primer lugar, esto sólo llevaría la pregunta a cómo sé que esa regla es válida, y en segundo lugar, sólo introduciría otra inferencia que tengo que hacer: Todas las inferencias de tal o cual forma son válidas; esta inferencia es de tal forma; por tanto, esta inferencia es válida". Por lo tanto, eso no ayuda. Tampoco hay que decir que los juicios lógicos en general se basan en argumentos, ya que entonces habría que averiguar la validez de estos últimos, lo que llevaría a un problema de circularidad o de regresión infinita.[31] Por último, los juicios lógicos tampoco se conocen por observación: la validez de un razonamiento no se ve con los ojos, ni se oye con los oídos, etc. Parece que la intuición es la única posibilidad que queda. Además, al hacer una introspección, nos damos cuenta de que tenemos intuiciones lógicas, y que éstas nos hacen pensar que algunas inferencias son válidas.

Este tipo de ejemplo es especialmente interesante, ya que todo razonamiento depende de los principios de la lógica. Por lo tanto, cualquier tipo de razonamiento depende de la intuición, incluyendo el razonamiento que el lector está haciendo en ese momento, e incluyendo cualquier razonamiento que pueda ser desplegado para impugnar la fiabilidad de la intuición.

Una posible respuesta a este argumento es que no necesitamos tener un conocimiento a priori de las verdades de la lógica, como por ejemplo que una determinada inferencia es válida; en su lugar, bastaría con que tuviéramos una disposición innata para hacer inferencias válidas. Aunque esta respuesta puede socavar la afirmación de que todo razonamiento depende de las intuiciones, no obvia la necesidad de la intuición en algún momento, por la sencilla razón de que, de hecho,

conocemos los principios de la lógica, y este conocimiento debe seguir siendo justificado. Entiendo que no se puede, sin alguna forma indeseable de circularidad, argumentar que una determinada forma de inferencia es válida utilizando un argumento de esa misma forma; por lo tanto, la cuestión sigue siendo que el conocimiento de las reglas de inferencia no puede ser en general inferencial.

Como en el ejemplo anterior, hay quien sostiene que las reglas de la lógica se hacen verdaderas "por definición" o por algún tipo de convenciones. La idea de que la verdad de las leyes de la lógica depende de las convenciones parecería sugerir que podríamos haber hecho convenciones o estipulaciones de tal manera que (sin cambiar el significado de ninguna de las siguientes palabras), la siguiente inferencia habría sido inválida:

Sócrates es un hombre.

Todos los hombres son inmortales.

Por lo tanto, Sócrates es mortal.

y la siguiente habría sido válida:

Sócrates es un hombre.

Por lo tanto, Sócrates es inmortal.

Por el bien de Sócrates, creo que deberíamos cambiar a convenciones de ese tipo.

Por supuesto, podríamos cambiar el uso de la palabra "válido" por convención. Pero eso es irrelevante; también podríamos cambiar el uso de la palabra "dientes" por convención, pero nadie considera que eso demuestre que el hecho de que los tiburones tengan dientes sea convencional en algún sentido relevante. (Si alguna vez te persigue un

tiburón, no te aconsejo que pongas tus esperanzas en un cambio oportuno de las convenciones lingüísticas). Cualquier afirmación verdadera podría convertirse en falsa con un cambio adecuado en el significado de las palabras que contiene. La cuestión es si el hecho que expresa el enunciado depende de una convención, es decir, si, una vez fijado el significado del enunciado, la convención desempeña algún papel más a la hora de determinar si lo que se dice es cierto. La prueba de ello es si podríamos hacer que el enunciado fuera falso mediante un cambio en alguna convención, sin cambiar lo que el enunciado significa.[32] Por ejemplo:

A. En Estados Unidos se conduce por el lado derecho de la carretera.

El enunciado (A) se hace verdadera por convención en un sentido sustantivo: hay una convención más allá de las que determinan el significado de las palabras de (A) que hace que (A) sea verdadera, a saber, nuestra convención de conducir por el lado derecho de la carretera. Si eliminá ramos esta convención, (A) se convertiría en falsa, sin ningún cambio en su significado. Ahora contrasta:

B. El silogismo, "Sócrates es un hombre; todos los hombres son machistas; por lo tanto, Sócrates es machista", es válido.

El enunciado (B) obviamente no depende de la convención en ese sentido. El significado de las palabras de (B) depende de las convenciones, como ocurre con todas las afirmaciones. Pero ninguna otra convención es relevante para la verdad de (B). No podemos hacer que (B) sea falsa cambiando ninguna convención, sin cambiar el significado de (B). Lo mismo ocurre con todos los principios lógicos. Las leyes de la lógica son, pues, ejemplos de hechos objetivos no convencionales que se conocen independientemente de la experiencia.

Esto es suficiente para dar una visión general de algunas de las dificultades para el empirismo. Otros han tratado esta cuestión de forma más completa y concluyente. Pero esto debería bastar para dejar claro por qué Mackie no tiene derecho a dar por sentado el empirismo como premisa para atacar el intuicionismo.

5.5 La inverosimilitud del nihilismo: un argumento mooreano

El nihilismo sostiene que nada es bueno, malo, correcto o incorrecto. Ya he dicho lo suficiente para mostrar por qué, a primera vista, está justificado rechazar esto. Un nihilista podría aceptar este punto, pero sostener que, no obstante, hay argumentos sólidos a favor del nihilismo que superan la presunción inicial en su contra.[33] En la última sección vimos algunas objeciones que un nihilista podría plantear contra el realismo, y veremos otras en capítulos posteriores. Lo que sostengo en esta sección es que la presunción contra el nihilismo es muy fuerte, de modo que los argumentos a favor del nihilismo tendrían que ser extremadamente poderosos para justificar la posición del nihilista.

Hasta ahora, me he centrado en el punto cualitativo de que muchas creencias morales tienen una justificación *prima facie*. Pero la justificación viene en grados: mi justificación para pensar que China existe es más fuerte que mi justificación para pensar que la teoría de la evolución es cierta, que es más fuerte que mi justificación para pensar que mañana hará sol. ¿Qué determina el grado de justificación *prima facie* de una creencia intuitiva? Si se acepta el conservadurismo fenoménico, la opinión natural es que cuanto más obvio parece algo, más fuerte es su justificación prima facie. Las intuiciones muy claras y firmes deben prevalecer sobre las débiles o vacilantes.

Consideremos ahora a grandes rasgos uno de los argumentos a favor del nihilismo:

1. El bien y el mal morales, si existen, serían intrínsecamente motivadores, es decir, cosas que cualquier ser racional estaría necesariamente motivado a perseguir (en el caso del bien) o evitar (en el caso del mal).

2. Es imposible que algo sea intrínsecamente motivador en ese sentido.

3. Por tanto, el bien y el mal no existen.[34]

Hay que decir más cosas para evaluar adecuadamente cada una de esas premisas, pero no lo diré ahora. Ahora mismo sólo quiero utilizar este argumento para ilustrar un punto epistemológico general. Dada la conclusión nihilista en (3), uno podría inferir válidamente otras conclusiones como:

4. No es cierto que una guerra nuclear sea mala.

5. Nunca se da el caso de que el disfrute sea mejor que el dolor insoportable.

Y así sucesivamente.

Ahora bien, así como alguien que aceptara (1) y (2) podría verse movido por el razonamiento anterior a aceptar (4) y (5), un realista podría argumentar contra (1) y (2) de la siguiente manera:

1". Una guerra nuclear sería mala.

2". El disfrute es a veces (si no siempre) mejor que el dolor insoportable.

3". Por lo tanto, el bien y el mal existen.

4". Por lo tanto, o bien:

a. El bien y el mal no tienen por qué ser intrínsecamente motivadores, o ...

b. Es posible que algo sea intrínsecamente motivador.

Algunos acusarían a este argumento realista de "plantear la cuestión" contra el nihilismo, ya que las premisas (1") y (2") son precisamente lo que el nihilista niega en su conclusión. Pero esto encarna una concepción ingenua de las cargas de la dialéctica, concediendo una presunción a cualquier argumento que se exponga primero. Porque si el argumento realista se hubiera enunciado primero, entonces podríamos decir que el argumento nihilista "se burla" del realista, ya que sus premisas (1) y (2) (conjuntamente) son precisamente lo que el realista niega en su conclusión. La relación entre los dos argumentos es simétrica: cada argumento toma como premisas la negación de la conclusión del otro argumento.[35] ¿Cómo, entonces, debemos decidir entre ellos?

La fuerza de un argumento depende de lo bien justificadas que estén las premisas y de lo bien que apoyen la conclusión. Los dos argumentos anteriores apoyan sus conclusiones igual de bien: ambos son deductivamente válidos. Por tanto, de los dos argumentos, el mejor es aquel cuyas premisas son inicialmente plausibles. Ahora bien, ¿cuál parece más obvio: "Disfrutar es mejor que sufrir un dolor insoportable" o "Es imposible que algo sea intrínsecamente motivador"? Para mí, lo primero parece mucho más obvio. Y no creo que mi juicio sobre este punto sea idiosincrático. Por lo tanto, sería irracional rechazar la primera proposición sobre la base de la segunda.[36]

Para justificar su posición, el nihilista tendría que presentar premisas más plausibles que cualquier juicio moral -más plausibles que "El asesinato está mal", más plausibles que "El dolor es peor que el placer", etc. Pero algunos juicios morales son tan plausibles como cualquier cosa. Por lo tanto, las perspectivas del nihilista parecen muy sombrías desde el

principio.

Por último, un comentario sobre el método filosófico. El argumento nihilista anterior, así como el argumento empirista discutido anteriormente (sección 5.4, Objeción 4), muestran un tipo de metodología racionalista común en la filosofía. El método es, a grandes rasgos, el siguiente: empezar por establecer como obvio algún principio abstracto de la forma "Ningún *A* puede ser *B*". (Por ejemplo, "Ningún conocimiento sustantivo puede ser a priori"; "Ninguna propiedad objetiva puede ser intrínsecamente motivadora"; "Ningún enunciado no verificable puede tener sentido"). A continuación, utilice el principio general como una restricción en la interpretación de los casos: si surgen casos de *A* que por todo el mundo parecen *B*, argumente que no pueden ser realmente *B* porque eso entra en conflicto con el principio, y busque alguna otra interpretación de los casos. Una de las grandes ironías de la filosofía es que esta metodología racionalista es empleada habitualmente por los empiristas. Cabría esperar que adoptaran el enfoque opuesto: empezar por observar los casos y sólo formar generalizaciones que se ajusten a la forma en que aparecen todos los casos; estar preparados para revisar cualquier generalización al descubrir contraejemplos; tratar los casos como una restricción de las generalizaciones.

Mi método se sitúa entre los dos anteriores: empezar con lo que parece cierto, tanto sobre los casos como sobre las reglas generales. Si surgen conflictos, los resuelvo a favor de la proposición que parezca más obvia. A grandes rasgos, queremos adoptar el sistema de creencias coherente que más se acerque a las apariencias, donde la fidelidad a las apariencias es una cuestión de cuántas proposiciones aparentemente verdaderas se mantienen, con estas proposiciones ponderadas por su grado inicial de plausibilidad. Podemos llamar a esto el método del equilibrio reflexivo.[37,38] El método del equilibrio reflexivo nos lleva a

aprobar algunos juicios morales. Es muy poco probable que pueda llevarnos a respaldar el nihilismo, ya que éste requiere un rechazo de todo nuestro conjunto de creencias morales. De hecho, sería difícil concebir una teoría menos fiel a las apariencias.

5.6 El realismo directo y la inversión subjetiva

Paso a otra objeción epistemológica al intuicionismo, que ayudará a aclarar el papel de la intuición en la producción de conocimiento. Consideremos un par de afirmaciones de la forma:

1. *S* tiene la intuición de que *p*.
2. *p*.

Donde *S* es una persona y *p* es una proposición. El intuicionista afirma que (1) desempeña algún papel en la justificación de (2) (al menos para *S*). ¿Pero cómo? No sólo (1) no implica lógicamente (2), sino que ni siquiera pertenecen a la misma materia. Supongamos que *p* es la proposición de que el camino más corto entre dos puntos es una línea recta. Entonces (2) es un enunciado geométrico; dice algo sobre la estructura del espacio. (1) no es un enunciado geométrico, sino psicológico; sólo dice que alguien tiene un cierto estado mental. ¿Qué relación puede tener (1) con (2)?

Las cosas estarían bien si pudiéramos argumentar que, a pesar de su independencia lógica, (1) hace que (2) sea más probable. Pero, ¿cómo podría la ocurrencia de un mero estado mental subjetivo hacer probable una proposición sobre el mundo externo? La sugerencia obvia es que lo haría si tuviéramos razones para creer que el estado mental es un signo del hecho externo; en otras palabras, que en general, cuando *S* tiene la intuición de que *p*, *p* suele ser verdadera. Para averiguarlo, parece que

tendríamos que observar una serie de casos en los que una persona tuviera una intuición, comparar esas intuiciones con los hechos y ver si, en general, existe una correlación. El problema en este proceso es la fase de comparación de las intuiciones con los hechos. Si, como parecen pensar los intuicionistas, no tenemos acceso a los hechos relevantes sin utilizar la intuición, entonces no se puede hacer tal comparación. Por lo tanto, es imposible decir si una intuición de que *p* es un signo de la verdad de *p* o no.

Esto, según mi oponente, agota las formas en que (1) podría ser relevante para la justificación de (2). El punto de preocuparse por la "justificación" de las creencias es que queremos que nuestras creencias sean en general verdaderas; una creencia justificada es una creencia fornida de tal manera que hace que sea al menos probable que sea verdadera.[39] (1) no garantiza la verdad de (2), ni tenemos ninguna razón para creer que aumenta la probabilidad de que (2) sea cierta. Así que aparentemente no tiene nada que ver con la justificación de (2).

Tal vez los que se resisten al intuicionismo tengan algo así en mente;[40] tal vez esto es lo que realmente subyace a las objeciones falaces de la sección 5.4. Los que rechazan este argumento se opondrían a la dependencia de las intuiciones en general, y no sólo a las intuiciones éticas.

Antes de discutir lo que está mal en este argumento, quiero hacer dos puntos preliminares para mostrar que es erróneo. El primero es que el argumento no es más que la adaptación a la intuición de un argumento clásico del escepticismo global.[41] Consideremos las afirmaciones:

3. Tengo una experiencia sensorial de *X*.[42]

4. *X* existe.

Donde *x* es algún tipo de cosa física. (3) no implica (4); (3) es un mero

informe psicológico, mientras que (4) es una afirmación sobre el mundo físico. Tampoco podemos verificar que (3) hace probable (4). Para ello, tendríamos que comparar la experiencia sensorial con la realidad física en varias ocasiones para establecer que una suele estar correlacionada con la otra de una manera determinada. No podemos realizar tal comparación, ya que no tenemos forma de acceder a la realidad física sin basarnos en la experiencia sensorial. Por lo tanto, parece que (3) no influye en la justificación de (4).

Del mismo modo, considere las afirmaciones:

5. Tengo una experiencia de memoria de *e*.

6. *e* ocurrió.

Donde *e* es un evento pasado putativo. Podemos argumentar que (5) no juega ningún papel en la justificación de (6). (5) se refiere a un estado mental presente, mientras que (6) se refiere a un acontecimiento pasado totalmente distinto (supuesto). Y no podemos confirmar que los recuerdos sean señales fiables de acontecimientos pasados, ya que no tenemos medios independientes de la memoria para acceder al pasado.

Creo que argumentos similares se aplican a todos los tipos posibles de conocimiento (supuesto), pero lo dejaré en esos dos ejemplos. Estos dos argumentos son exactamente análogos al argumento escéptico sobre la intuición. Dado que asumo que no queremos apoyar ni el escepticismo sobre el mundo físico ni el escepticismo sobre el pasado, debemos buscar el error en el argumento.

El segundo punto destinado a demostrar que el argumento debe ser erróneo es su carácter autorrefutatorio. Las intuiciones no son entidades teóricas exóticas inventadas por algunos filósofos. No desempeñan simplemente un papel menor de investigación, de modo que podríamos suprimirlas y nuestra vida intelectual seguiría más o menos como hasta

ahora. Tampoco existe una metodología alternativa, independiente de las intuiciones, que esté siendo aplicada por algún otro grupo de filósofos. Las intuiciones no son más que apariencias intelectuales iniciales. Es decir, son la forma en que las cosas parecen, intelectualmente, antes de la argumentación. Comprenden cosas como (i) la "verosimilitud" de las premisas de partida de los argumentos filosóficos, (ii) nuestra "visión" de la conexión lógica entre las premisas y la conclusión de un argumento, (iii) los juicios generales de verosimilitud que recorren el discurso y el razonamiento intelectuales, ya sea de forma explícita o no. En el argumento escéptico con el que he abierto esta sección, ¿cómo es que lo sabemos?

> Que (1) ("*S* tiene una intuición de que *p*") no entraña (2) ("*p*");
>
> Que los enunciados psicológicos no conllevan los geométricos;
>
> Que (1) hace probable (2) sólo si (1) es un signo de (2);
>
> Que (1) es un signo de (2) sólo si las intuiciones en general están correlacionadas con los hechos;
>
> Que no podemos verificar que las intuiciones están correlacionadas con los hechos a menos que tengamos un conocimiento no intuitivo de los hechos;
>
> Que (1) desempeña un papel en la justificación de (2) sólo si (1) implica (2) o hace que (2) sea probable?

Por supuesto, intentar dar argumentos para estas cosas me llevaría simplemente a preguntar cómo conocemos las premisas de esos argumentos. Cuando usted leyó el argumento escéptico tal como lo presenté inicialmente -si el argumento le pareció razonable- estaba "intuyendo" todas esas cosas. Las intuiciones no son algo que sólo utilicen unos pocos filósofos, los intuicionistas; otros filósofos simplemente se abstienen de llamar a sus intuiciones "intuiciones". En

consecuencia, la idea de que hay dos formas de proceder en el discurso intelectual -apelar a las intuiciones y "demostrar que las afirmaciones propias son verdaderas"- es pura confusión. Todo lo que puedo hacer para "demostrar que algo es cierto" en un artículo o un libro es escribir una serie de frases. Las frases cuentan como demostración de algo sólo si, al leerlas, parecen verdaderas (al menos algunas de ellas) y parecen apoyar mi conclusión.[43] La apelación a las intuiciones es, por tanto, una parte integral de la demostración de la verdad, no una alternativa a la demostración de la verdad.

Las observaciones anteriores pretenden motivar la búsqueda de un fallo en el argumento escéptico y hacernos receptivos a los diagnósticos propuestos. No explican cuál es ese defecto. En mi opinión, el defecto consiste en un malentendido básico de la estructura de una teoría fundacionalista del conocimiento. El intuicionismo no sostiene que de "tengo una intuición de que *p*" se pueda inferir "*p*"; ni el principio del conservadurismo fenomenal sostiene que "me parece que p" es una razón para "*p*". Esas serían afirmaciones sobre la justificación inferencial.[44] El conservadurismo fenomenal y mi versión de la intuición son formas de *fundacionalismo*: sostienen que estamos justificados en algunas creencias sin necesidad de pruebas. El papel de las condiciones (1), (3) y (5) en la teoría de la justificación es el de las condiciones bajo las cuales ciertas creencias -respectivamente, las expresadas por (2), (4) y (6)- no requieren evidencia, y no el de la evidencia que apoya esas creencias.[45] El punto lógico de que (1), (3) y (5) no proporcionan argumentos convincentes para (2), (4) y (6) es, por tanto, irrelevante para la teoría.

Los escépticos no están de acuerdo; los escépticos globales negarán que exista alguna condición que permita justificar una creencia en ausencia de razones para ello. Pero no escribo ahora para convencer a los escépticos globales. Escribo para los no escépticos, para diagnosticar el

escepticismo. El escéptico, sugiero, tiene una teoría sobre la relación entre la conciencia y sus objetos, según la cual aquello de lo que somos inmediatamente conscientes debe ser siempre un estado interno, mental. Desde este punto de vista, cuando percibimos, somos conscientes, ante todo, de las experiencias sensoriales; cuando recordamos, somos conscientes de las imágenes de la memoria; cuando contemplamos asuntos abstractos, somos conscientes de los conceptos y las intuiciones. La función de las experiencias sensoriales, las experiencias de memoria y las intuiciones es, por tanto, meramente la de piezas de evidencia a partir de las cuales intentamos inferir conclusiones sobre la realidad extra-mental. El fracaso de esta inferencia, a su vez, es una victoria para el escepticismo.

Si aceptamos el punto de partida del escéptico, creo que el resto de su razonamiento es convincente. Pero ese punto de partida no es natural ni está respaldado por ningún buen argumento. Un punto de vista más natural es que, en primer lugar, somos conscientes de las cosas, es decir, de las cosas externas. A continuación, reflexionamos sobre nuestra conciencia de esas cosas, tras lo cual nos damos cuenta (o tal vez inferimos) que esa conciencia implica que tenemos estados internos que de alguna manera representan las cosas externas. Estos estados internos no deben suplantar a los objetos reales en nuestra filosofía; su función central es la de vehículos de la conciencia de las cosas externas. La "inversión subjetiva" del escéptico convierte los estados cognitivos que se plantearon para explicar nuestra conciencia del mundo en un velo que bloquea nuestra visión del mundo, y hace que la conciencia sea esencialmente introvertida, contemplándose sólo a sí misma. Por el contrario, el punto de vista más natural, el "realista directo", es que la función principal de la experiencia sensorial es constituir en parte nuestra conciencia de las cosas externas, en lugar de ser un objeto interno1ediario de la conciencia. Una idea relacionada con la filosofía de

la percepción, denominada "la transparencia de la experiencia", sostiene que el modo en que determinamos las propiedades de nuestras experiencias sensoriales es mirando los objetos que percibimos; cuando intentamos mirar nuestras experiencias, simplemente "vemos a través de ellas" los objetos que representan.[46]

Esta es una declaración aproximada y resumida del realismo directo en la filosofía de la percepción. ¿Qué tiene que ver con la filosofía moral? El intuicionista, sostengo, debería ser un realista directo sobre la ética. No debería decir que las intuiciones funcionan como una especie de evidencia de la que hacemos o debemos inferir conclusiones morales. Debería decir que, para algunas verdades morales, no necesitamos pruebas, ya que somos directamente conscientes de ellas, y esa conciencia toma la forma de intuiciones; es decir, las intuiciones sólo constituyen (en parte) nuestra conciencia de los hechos morales.[47] Las intuiciones no son los objetos de nuestra conciencia cuando hacemos filosofía moral; son sólo los vehículos de nuestra conciencia, a través de los cuales "vemos" la realidad moral.

Esto no implica que la intuición sea infalible. Algunas intuiciones pueden ser erróneas, en cuyo caso no constituyen una conciencia directa de los hechos morales. Pero esto no impide que las restantes intuiciones, las que son verdaderas, constituyan una conciencia directa de los hechos morales. (Compárese: las alucinaciones no constituyen conciencia de los objetos físicos, pero esto no impide que la percepción normal constituya tal conciencia).

La tesis de la transparencia de la percepción, al igual que la tesis paralela de la transparencia de la intuición, es una tesis fenomenológica: en la percepción, se nos presentan fenómenos físicos de diversos tipos; no se nos presentan estados mentales. Incluso si la presentación fuera falsa (como en una alucinación), seguiríamos sin ser conscientes de un

estado mental; entonces simplemente no seríamos conscientes de nada real, aunque pareciera como si lo fuéramos. Del mismo modo, en la intuición ética, como punto de hecho fenomenológico, nos encontramos con propiedades y relaciones morales, no con estados mentales. Esto bloquea el intento de interpretar las experiencias y las intuiciones como pruebas a partir de las cuales hacemos inferencias (falaces, como se vería después) sobre el mundo.

*5.7 El aislamiento del realismo moral

Concluyo con una última objeción epistemológica al intuicionismo. Incluso si las propiedades morales son reales, no parece que puedan afectar a nada. No producen efectos físicos, por lo que no afectan a nuestros procesos cerebrales, así que probablemente tampoco afecten a nuestros procesos mentales.[48] He aquí un argumento para ello: Las verdades evaluativas básicas que intuimos son verdades necesarias, como "El disfrute es intrínsecamente bueno".[49] Estos hechos son atemporales; es decir, no hay cambios en el ámbito de las verdades morales fundamentales. Un acontecimiento concreto puede convertirse en bueno o dejar de serlo, por ejemplo, convirtiéndose en placentero o dejando de serlo; pero el hecho abstracto de que el disfrute es bueno no puede dejar de existir. Pero esto significa que nunca ocurre nada -no hay acontecimientos- en el ámbito puramente moral. Y eso implica que el ámbito moral es causalmente inerte.

Algunos filósofos sostienen que el conocimiento de una cosa requiere algún tipo de interacción con ella.[50] Por ejemplo, puedo conocer al gato blanco del sofá porque su presencia produce un efecto en mí (me hace tener una experiencia sensorial que representa al gato blanco). Pero los hechos morales no tienen ningún efecto sobre mí; por tanto, no conozco

ningún hecho moral.

Un argumento estrechamente relacionado parte de la premisa de que conozco *p* sólo si, de no ser cierto *P*, no creería que lo fuera.[51] Pero si los hechos morales fueran distintos de como son en realidad, o si no hubiera hechos morales, tendría las mismas creencias morales que tengo actualmente, puesto que los hechos morales no causan mis creencias. Por tanto, no conozco los hechos morales.

Es bastante fácil responder rechazando los relatos de conocimiento asumidos en los dos párrafos anteriores. Hay serias objeciones a ambas -entre las que destaca el hecho de que no pueden dar cuenta del conocimiento a priori- y hay varios análisis alternativos del "conocimiento" disponibles.[52]

Pero no digo que no haya ningún problema aquí. Hay una condición más general sobre el conocimiento que todo el mundo acepta en epistemología: Conozco *p* sólo si no es un mero accidente (no una cuestión de azar) que esté en lo cierto sobre si *p*. El reto para el realista moral, entonces, es explicar cómo sería algo más que un azar si mis creencias morales fueran verdaderas, dado que no interactúo con las propiedades morales.

Consideremos una respuesta abortada: No es casualidad que mis creencias morales sean verdaderas, porque los hechos morales, en general, son necesarios; no podrían haber sido de otra manera... Por ejemplo, no hay ningún mundo posible en el que el disfrute no sea intrínsecamente bueno. Por tanto, no es casualidad que tenga razón cuando afirmo que el disfrute es bueno.

Para ver por qué esta respuesta es errónea, imaginemos a alguien que comete un error de cálculo al multiplicar 57 por 69, pero luego comete un segundo error que casualmente anula el primero, de modo que llega a

la respuesta correcta (3933) de todos modos. En este caso, lo que cree es necesariamente cierto: no hay ningún mundo posible en el que 57 x 69 no sea igual a 3933. Pero sigue siendo accidental que su creencia sea verdadera. ¿Por qué? Porque normalmente no se puede esperar que el hecho de cometer uno o dos errores en un cálculo haga que se llegue a la conclusión correcta. Cuando preguntamos por qué no es accidental que una creencia sea verdadera, lo que queremos decir es: ¿cómo la forma en que se formó la creencia hace predecible que sea verdadera?

Ahora bien, este problema no es específico del conocimiento moral. Es un problema general sobre el conocimiento a priori. Paul Benacerraf lo planteó originalmente como un problema sobre las matemáticas: puesto que no tenemos ninguna interacción con el número 2 -no nos topamos con él en la calle, etc.- ¿cómo podemos tener conocimiento de él? Podría alegar que no es tarea del filósofo moral responder a esto. Haya o no conocimiento moral, hay conocimiento a priori de otros tipos, por lo que debe haber alguna solución al problema de Benacerraf. Cualquiera que sea la explicación para el conocimiento a priori en general, no hay razón para pensar que no funcionaría igual de bien para el conocimiento moral.

Creo que esta sería una respuesta justa. Sin embargo, si me quedara con esto, algunos lectores se sentirían decepcionados y querrían conocer mi punto de vista sobre el conocimiento a priori. Como éste no es un libro ni de epistemología ni de metafísica, sólo daré un breve resumen. Mi visión del conocimiento a priori tiene cuatro elementos principales.

Primero: Los universales existen necesariamente. Los "universales" son cosas abstractas (características, relaciones, tipos) que dos o más cosas o grupos particulares pueden tener en común. Por ejemplo, el amarillo es un universal. Es algo que tienen en común los limones, el sol y los autobuses escolares, entre otras cosas. El amarillo es "abstracto" en

el sentido de que no es un objeto concreto con una ubicación concreta; no te vas a encontrar con un amarillo solo por la calle. Sin embargo, el amarillo existe. He aquí un argumento a favor:

1. La siguiente afirmación es verdadera: (Y) El amarillo es un color.

2. La verdad de (Y) requiere que el amarillo exista.

3. Por lo tanto, el amarillo existe.[53]

Comentario: Supongamos que digo: "El rey del Colorado es esponjoso". Como no existe el rey de Colorado, algunos dirían que la frase es falsa; otros dirían que no es ni verdadera ni falsa. Pero nadie cree que sea verdadera. Del mismo modo, supongamos que digo: "Dios es amoroso". Obviamente, esta frase no podría ser verdadera a menos que "Dios" se refiera a algo, es decir, a menos que Dios exista. La oración (Y) tiene la misma forma, por lo que sólo puede ser verdadera si "amarillo" se refiere a algo, es decir, sólo si el amarillo existe.

Algunos filósofos (los "nominalistas") dicen que lo único que tienen en común los particulares múltiples es que les aplicamos la misma palabra o idea.[54] He aquí un argumento en contra:

4. El amarillo es un color, y los limones lo tienen.

5. Ninguna palabra o idea es un color, ni los limones "tienen" palabras o ideas.

6. Por lo tanto, el amarillo no es una palabra ni una idea.

El amarillo es algo que tienen en común los limones, el sol, etc.; por tanto, lo que tienen en común no es (meramente) una palabra o una idea. Algunos filósofos dirán que he simplificado demasiado esta cuestión. Yo digo que he simplificado pero no simplificado en exceso; la existencia de universales es una verdad trivial.

Segundo: Al tener conceptos, captamos (entendemos) los universales.[55] Tener el concepto de amarillo es comprender (al menos en parte) lo que es el amarillo. La captación se produce en grados: uno puede captar algo mejor o peor. Una comprensión adecuada de un universal es un concepto que es:

i) Consistente. Por ejemplo, el concepto de cuadrado redondo es inconsistente; en consecuencia, no cuenta como una comprensión adecuada de un universal. Existen incoherencias más sutiles, como la que puede suponer el concepto del mayor número primo.

ii) Claro, en contraposición a confuso. Por ejemplo, alguien que haya leído un poco sobre la teoría del caos puede pensar que el "caos" es como el "azar". Es decir, puede no distinguir estos dos conceptos. En ese caso, su concepto de caos es confuso y no es una comprensión adecuada de la naturaleza del caos.

iii) Determinado, por oposición a vago o inestable. Imagine que alguien discute sobre el aborto. Le hablas de la píldora RU-486 y le preguntas si cuenta como aborto. Él no lo sabe. En este caso, su concepto de aborto es indeterminado; los criterios para aplicar el concepto no están establecidos en su mente.

Las características anteriores se dan en grados; decimos que la comprensión de una persona es adecuada si su concepto tiene estos rasgos en un alto grado.[56]

Tercero: Tener una comprensión adecuada de un universal lo pone a uno en posición de ver que tiene ciertas propiedades y/o relaciones con otros universales que usted capta adecuadamente. Si uno sabe lo que es el amarillo (capta adecuadamente la amarillez), puede intuir que el amarillo es un color. Si sabes lo que son los números dos y tres (captas adecuadamente el dos y el tres), entonces puedes intuir que el dos es menos que el tres. Cualquiera que no pudiera ver estas cosas tendría que

haber fallado en la comprensión de los universales relevantes. Esto no se debe a que estas cosas sean "analíticas", o verdaderas en virtud de las definiciones de los conceptos relevantes, ya que, en mi opinión, casi ningún concepto es definible (en el sentido de los filósofos...); se debe a que la comprensión de la naturaleza de un universal tiende inherentemente a hacer que uno aprehenda ciertos hechos básicos sobre él.[47] Por supuesto, no todas las intuiciones son esenciales para la comprensión de los conceptos relevantes, pero una persona que tenga alguna comprensión de un universal tendrá necesariamente al menos algunas intuiciones correctas sobre él.

Los fundamentos de la segunda y tercera afirmaciones anteriores son introspectivos: la introspección revela que a veces comprendemos conceptos, y que a veces tenemos la experiencia de ver que algo es verdadero, debido a esa comprensión.

Cuarto: Todo conocimiento a priori es, o se deriva, del conocimiento de las propiedades y relaciones de los universales. Por ejemplo, sabemos a priori que todas las solteronas son solteras. Esto se deriva de nuestro conocimiento de que la soltería contiene o implica la soltería. Sabemos a priori que el púrpura es un color: esto es el conocimiento de una propiedad del púrpura. Y sabemos a priori que el placer es bueno: esto es el conocimiento de una propiedad del placer.

Ahora, volviendo al problema de Benacerraf: la cuestión es, dado el modo en que se forman (al menos a veces) las creencias a priori, ¿por qué es más que accidental que sean verdaderas? Para empezar, propongo que tener un concepto claro, consistente y determinado es suficiente para que uno capte un universal o universales. No hay posibilidad de que uno no se refiera a nada (los universales son abundantes en este sentido, y su existencia es necesaria).[58] Nótese, sin embargo, que las características definitorias de una captación adecuada son intrínsecas: la consistencia,

claridad y determinación pertenecen a la naturaleza de un concepto en sí mismo, en oposición a depender de las relaciones entre el concepto y algo más. Así pues, las características intrínsecas de un concepto a veces son suficientes para que constituya una comprensión adecuada de la naturaleza de un universal. Además, la comprensión adecuada de un universal no puede causar intuiciones falsas sobre él.[59] Por lo tanto, en algunos casos -a saber, cuando las intuiciones de uno están causadas (únicamente) por una comprensión clara, coherente y determinada- el proceso interno por el que uno forma creencias garantiza su verdad.

Algunos dirán que "garantiza" es demasiado fuerte; basta con decir "hace altamente probable". Pero creo que la afirmación "garantiza" es correcta. Obsérvese que no se afirma que todas las intuiciones sean verdaderas. Tampoco se trata de que todas las intuiciones de una persona que comprenda adecuadamente los conceptos relevantes sean verdaderas. Como veremos en el capítulo 6, podemos equivocarnos de muchas maneras. La afirmación es que captar adecuadamente un concepto no puede causar un error por sí mismo. En otras palabras, si una persona tiene una intuición falsa, esta intuición falsa debe estar causada por otra cosa: un malentendido, un prejuicio, una confusión o algo parecido. Es coherente con esto que haya muchas intuiciones falsas. Esto no socavaría mi pretensión de resolver el problema de Benacerraf. El problema consistía en explicar cómo la forma en que se forman algunas creencias a priori hace predecible que sean verdaderas. Las creencias que se forman mediante procesos que implican sesgo, confusión, etc., se forman de un modo distinto a las creencias que son causadas por una comprensión adecuada sin tal sesgo, confusión, etc. Por lo tanto, la falta de fiabilidad del primer tipo de creencias no impide que el segundo tipo de creencias se considere conocimiento.

Lo anterior no es una teoría de la justificación de las creencias a

priori. Mi opinión no es que se justifiquen las creencias a priori mediante las reflexiones que acabo de hacer. Más bien, mi teoría de la justificación se dio en las secciones 5.1-5.2. Lo anterior se dirigía a un problema distinto, el problema de por qué, aunque fueran verdaderas y estuvieran justificadas, nuestras creencias morales no serían meramente verdaderas por accidente.[60]

Notas

1. Véase mi (2001, pp. 98-115) y (de próxima publicación). Véase Bonjour (2004) para las objeciones al principio. El término inglés "phenomenon" deriva del griego plwillumenoll, que significa "apariencia"; de ahí que el conservadurismo fenomenal sea un conservadurismo con respecto a las apariencias. Es decir, sostiene que debemos evitar apartarnos de las apariencias innecesariamente.

2. La tercera es un poco extraña; normalmente, uno se limitaría a informar de que le duele la cabeza. Informar de que uno parece tener dolor de cabeza es extraño debido a la implicación conversacional de que uno está en posición de no decir nada más fuerte. Sin embargo, aunque sea extraño, creo que el informe "parece" sería verdadero. Esto no tiene por qué significar que la apariencia de que uno tiene dolor sea un estado mental separado del dolor; quizás, más bien, algunos estados mentales tienen conciencia de sí mismos incorporada.

3. Una persona puede creer lo que no parece cierto ejerciendo el autoengaño o dando un salto de fe; véase mi (2001, pp. 108-10). Pero estas formas de creer son epistémicamente irracionales.

4. Compárese la concepción de la intuición de Bealer (1992, pp. 101-2). Esta no es la única concepción de la intuición; en contraste con Bealer y conmigo, algunos filósofos ven las intuiciones como una especie de creencia (Audi 2004, pp. 33-6). Nótese que cuento las inferencias inconscientes e implícitas como una especie de "razonamiento", aunque no contaría el procesamiento de información que es en principio inaccesible a la reflexión, como el procesamiento preconsciente realizado por el sistema visual, como una forma de "razonamiento".

5. Fumerton (1990, p. 6) no tiene "ni idea de lo que se supone que son estas

intuiciones". Ayer (1956, p. 31) dice que decir que alguien sabía algo por intuición "no es más que afirmar que lo sabía pero que no podemos decir cómo". Véase también la nota 8 más abajo.

6. Tara Smith (2000, p. 24).

7. Si los dos enunciados "si... entonces" deberían contar como enunciados evaluativos es un punto discutible, pero meramente semántico; en cualquier caso, serían de utilidad en el razonamiento posterior de la ética normativa.

8. Hudson (1967, p. 57; 1983, pp. 103-4); Brandt (1979, pp. 17,20); Bennett (1995, p. 12).

9. El ejemplo 1 procede originalmente de James Rachels (en conversaciones informales de los años sesenta) y se desmenuza más en Harman (1977, pp. 3-4). El ejemplo 2 procede de Foot (1967).

10. Compárese Smart (1973, pp. 69-72), que se enfrenta a un ejemplo similar y admite que algunas consecuencias de su utilitarismo son muy desagradables.

11. Popper (2001, pp. 49-50). La observación "lo que hemos visto que es verdadero debe serlo" es engañosa; Popper no podría objetar esto, ya que es una tautología, puesto que "ver" es factivo. Más bien, lo que pone en duda es que todas las intuiciones cuenten como ver que algo es verdadero.

12. Entre los intuicionistas, Moore (1960, p. x), Audi (1997, pp. 46, 55; 2004, pp. 30-2), Rashdall (1907, p. 85) y Lemos (1994, pp. 144-52) rechazan explícitamente la infalibilidad. Ross afirma la certeza de algunos principios morales (1988, pp. 29-30), pero a continuación niega que los juicios sobre actos concretos sean siempre seguros (pp. 30-1) y más adelante permite que las convicciones morales a veces deban corregirse (p. 41). Sterling (1994, pp. 130-1) afirma que las intuiciones son necesariamente verdaderas, ya que son una especie de conocimiento directo, pero que, sin embargo, todas las creencias éticas son falibles. Prichard llama a los principios éticos "autoevidentes", pero lo define sólo en el sentido de que no son derivados (1957, p. 8); admite la posibilidad de fallos en la percepción de las verdades morales (pp. 9-IOn). Shafer-Landau considera que algunos principios morales son irrevisables, pero niega que esto sea necesario para que un principio cuente como 'autoevidente' (2003, p. 249).

13. Sinnott-Armstrong (1996, p. 10).

14. Stratton-Lake (2002a, p. xiii); Shafer-Landau (2003, pp. 247-50); Sterling (1994, p.

128).

15. Estos malentendidos aparecen en Tara Smith (2000, pp. 25-6), quien escribe: "La afirmación de los intuicionistas sobre la irrelevancia de los argumentos es plausible sólo en la medida en que los ejemplos invocados sean cuidadosamente seleccionados", concediendo aparentemente que algunos conocimientos morales son plausibles de ser considerados intuitivos.

16. Pero véase más arriba, nota 12. No encuentro ningún intuicionista que afirme la infalibilidad moral.

17. El pasaje aparece en "Does Moral Philosophy Rest on a Mistake?" (Prichard 1957, p. 16), que, a pesar de algunas afirmaciones exageradas, es uno de los trabajos más relevantes en este campo. Nótese que, a pesar de la afirmación de "autoevidencia" que he citado, Prichard sostiene que pueden ser necesarios ciertos "preliminares" para ver que uno tiene una obligación particular en un caso concreto, lo que puede incluir la recopilación adicional de información no moral (pp. 7-8).

18. La idea de que el fundacionalismo requiere incorregibilidad es refutada por Audi (1983). Alston (1976) también refuta las objeciones al fundacionalismo que dependen de este malentendido.

19. Tara Smith (2000, p. 24); Brandt (1996, p. 5).

20. Alston (1993) sostiene que es imposible dar un argumento no circular para la fiabilidad de la percepción de los sentidos. Véase especialmente (1993, pp. 115-19), donde aplica el argumento a todas las formas de formar creencias.

21. Para las defensas del coherentismo, véase Davidson (1990) y Bonjour (1985), pero nótese que Bonjour acepta el coherentismo sólo para el conocimiento empírico, no para el conocimiento a priori (p. 193). Para las objeciones, véase Foley y Fumerton (1985); Huemer (1997); y Fumerton (1995, pp. 144-7).

22. Hudson (1983, pp. 104-5); Ayer (1952, pp. 106, 118-20).

23. Sterling (1994, p. 79) hace referencia a la introspección. Véase mi (1999, pp. 347-8) sobre el problema de la memoria. Véase Hume ([1758] 1975, pp. 32-9) sobre el problema de la inducción.

24. Thomson (1971).

25. Sinnott-Armstrong (1996, p. 10). Sextus Ernpiricus (1964, p. 74) formula la misma

acusación contra el fundacionalismo en general.

26. Mackie (1977, pp. 38, 39).

27. Como observa Sterling (1994, p. 67).

28. Véase también mi discusión sobre el argumento de la rareza en la sección 8.1 más adelante.

29. Véase el excelente III Defense of Pure Reason de Bonjour (1998) y Bealer (1992).

30. Crane y Pian tanida (1983) afirman haber producido experiencias sensoriales de verde rojizo, o de un color que parece estar entre el rojo y el verde. Algunos filósofos creen que esto demuestra que se puede tener una experiencia de que algo es totalmente rojo y totalmente verde. Para estos filósofos,l observar que "*x* es rojizo" no implica "*x* es rojo", ni "*x* está entre *A* y *B*" implica "*x* es tanto *A* como *B*".

31. Esencialmente este punto aparece en Russell ([1912] 1997, pp. 72-4) y Bonjour (1998, pp. 43-4).

32. Como apunte, este criterio implica que las típicas verdades "analíticas" como "Todos los dodecaedros tienen 12 caras" no dependen de la convención en ningún sentido interesante. Creo que esto es correcto. El hecho de que todos los dodecaedros tienen 12 caras se mantendría incluso si no tuviéramos lenguaje, aunque la frase "Todos los dodecaedros tienen 12 caras" obviamente no. Véase Quine (1976) para las críticas relacionadas con la noción de verdad por convención.

33. Esto es justo lo que hace Mackie (1977, p. 35).

34. Mackie (1977, pp. 35, 40); joyce (2001, pp. 17-27).

35. Nelson (2003, pp. 76-7) hace una observación similar, inspirada en las observaciones de Chisholm (1982) sobre el problema del criterio, pero Nelson concede que ambos argumentos plantean la cuestión. Mi opinión es que ninguno de los dos lo hace (en la medida en que plantear la cuestión es una falacia). Chisholm también sostuvo que su propia solución al problema del criterio planteaba una falacia, pero lo disculpó con la observación de que todas las soluciones plantean una falacia. Una vez más, yo negaría que cualquiera de las posiciones que él discute (particularismo, metodismo y escepticismo) se haga la pregunta. Véanse mis observaciones sobre la evasión de la cuestión en la sección 4.2 anterior. James Rachels (1974, págs. 310-13) hace observaciones similares sobre la "evasión de la pregunta" contra el egoísmo ético.

36. El argumento aquí es análogo a los que se encuentran en Moore (1953, pp. 119-20), DePaul (1988, p. 85) y Huemer (2001, pp. 31-44). En este último, sostengo que toda forma de escepticismo filosófico es injustificada; el argumento presentado aquí es un caso especial de ese argumento.

37. Compárese la discusión de Chisholm (1982) sobre el metodismo y el particularismo, aunque Chisholm descuida el enfoque del equilibrio reflexivo. El término "equilibrio reflexivo" procede de Rawls (1971, p. 48). DePaul (1988) defiende este enfoque general de la teoría moral, al que llama "coherentismo". En mi opinión, el enfoque no se llama propiamente coherentista, ya que hay que conceder algún grado de justificación fundacional a las apariencias iniciales. Véase mi (1997) y Olsson (2005) para la justificación de este requisito.

38. Sin embargo, esto es una especie de simplificación de mi punto de vista, ya que implica que la "coherencia" es una propiedad de todo o nada. Una afirmación más precisa incorporaría la noción de grados de coherencia, donde éstos están parcialmente determinados por las relaciones probabilísticas entre las creencias. Además, la afirmación del texto implica que lo que importa es el número absoluto de creencias inicialmente plausibles que se mantienen. Una afirmación más precisa tendría en cuenta el hecho de que algunos pares de creencias inicialmente plausibles están tan relacionados que, si uno de ellos es falso, el otro es casi seguro que también es falso, de modo que renunciar a la última creencia no debería contarse como un coste adicional de una teoría que requiere renunciar a la primera creencia. Pero no creo que estos refinamientos sean importantes para el presente propósito, simplemente porque el nihilismo en su cara está muy lejos de lograr cualquier variante razonable de equilibrio reflexivo.

39. Esta es una concepción popular de la justificación epistémica; véase Bonjour (1985, pp.7-8).

40. Compárese Hare (1989, pp. 82-3).

41. Seemy (2001, pp. 1l-16, 178-81); Hume ([1758] 1975, pp. 151-3). El escepticismo global es la opinión de que nadie puede estar justificado en la creencia de nada.

42. Una experiencia sensorial de *x* es, a grandes rasgos, la experiencia de parecer percibir x; nótese que el hecho de tener tal experiencia no implica que *x* exista realmente. La "experiencia de memoria" debe entenderse de forma análoga.

43. Que yo demuestre que algo es verdadero es, por supuesto, compatible con que algunos lectores no lo vean. Pero si, al leer mi trabajo con atención, la mayoría de los

jueces competentes no encontraran que ninguna de mis premisas pareciera verdadera, entonces, incluso si todas mis afirmaciones fueran de hecho verdaderas, no contaría como si hubiera demostrado algo.

44. Nagel (1986, p. 146) parece respaldar esta concepción inferencial del papel de la intuición ética: "[E]n la ética, uno infiere desde las apariencias de valor hasta su explicación más plausible en una teoría de lo que hay razón para hacer o querer". Esta observación es, sin embargo, ambigua; "apariencia" puede referirse tanto al estado mental de algo que parece ser el caso como al contenido de ese estado mental, lo que parece ser el caso. En una interpretación, Nagel nos aconseja adoptar la mejor explicación de por qué ciertas cosas parecen ser valiosas. En la otra interpretación, nos aconseja adoptar la mejor explicación de por qué ciertas cosas son valiosas, cuando éstas son las que nos parecen valiosas. Es la primera doctrina la que objeto como un malentendido del papel epistemológico de la intuición.

45. Obsérvese que (1) puede ser una razón por la que la creencia en (2) está justificada, en el sentido de una condición a la que sobreviene la justificación para creer en (2), sin ser la razón de o/le o incluso una razón que olle tiene para creer en (2), en el sentido de algo de lo que uno hace o podría inferir legítimamente (2). Véase mi (2003, pp. 148-51) sobre esta distinción.

46. Harman (1990) y Tye (2002) defienden la tesis de la transparencia. Reid (1983) ofrece una forma temprana de realismo directo. Elaboro y defiendo este punto de vista en mi (2001). Véanse especialmente las pp. 81-5 sobre la concepción errónea del escéptico de la relación entre la conciencia y sus objetos.

47. Véase mi (2001, pp. 5l-5) sobre la naturaleza de la conciencia. A grandes rasgos, las intuiciones (a veces) cuentan como conciencia de los hechos morales porque (a veces) los hechos morales corresponden a los contenidos proposicionales de las intuiciones, las in tuiciones representan sus contenidos como reales, y no es cierto que estas intuiciones sean verdaderas.

48. Esto se deduce si se acepta el fisicalismo. Pero incluso los dualistas mente/cuerpo probablemente aceptarán que las propiedades morales no causan eventos mentales. Obsérvese que todos los argumentos de esta sección deben considerarse relativos a hechos morales fundamentales, de la forma "*F* es hacer el bien", "*G* es hacer el mal", etc.

49. De hecho, preferiría considerarlos como principios sobre lo que son rasgos de hacer el bien, hacer el bien, etc.; véase más adelante, sección 8.2.3.

50. Goldman (1967); Benacerraf (1973, p. 671). La presente objeción al conocimiento moral está modelada sobre la objeción de Benacerraf al platonismo en las matemáticas.

51. Véase el análisis del conocimiento de Nozick (1981, pp. 172 y ss.).

52. Las objeciones que tengo en mente incluyen la dificultad de la teoría causal para dar cuenta del conocimiento por inducción, y los fallos de cierre que implica la teoría de Nozick. Los relatos alternativos del conocimiento que tengo en mente incluyen el confiabilismo, el análisis de la derrotabilidad, el análisis de la "función adecuada" de Plantinga, la teoría de las alternativas relevantes y varias formas de contextualismo.

53. Véase Armstrong (1978a; 1978b) para más argumentos en defensa de los universales y (1978a, pp. 60-1) sobre la dificultad que oraciones como (Y) plantean a los nominalistas; pero nótese que Armstrong acepta muchos menos universales que yo. Bealer (1993) proporciona argumentos en apoyo de una premisa análoga a mi (2), relativa a los enunciados sobre las propiedades de las proposiciones. Pero dudo que ninguno de estos filósofos esté de acuerdo conmigo en que la existencia de universales es una verdad trivial.

54. Berkeley ([1710]1965, introducción, secciones 11-16); Hume ([1739J 1992, pp. 17-25). Hay otras formas de nominalismo que no discuto aquí, particularmente lo que Armstrong llama "nominalismo de clase" y "nominalismo de semejanza". Creo que pueden ser refutados de forma similar; véase Armstrong (1978a, capítulos 4, 5).

55. Bealer (1999) y Russell ([1912] 1997, capítulos 9-10) ofrecen puntos de vista similares sobre el conocimiento a priori. Considero que la noción de Russell de tener "conocimiento" de los universales y la noción de Bealer de poseer un concepto "en el sentido pleno" son esencialmente equivalentes a mi noción de "captar adecuadamente" un universal. Lemos (1994, pp. 155-60) y Audi (1997, p. 28; 2004, p. 48) aplican una concepción similar del conocimiento a priori al caso de la ética, pero observan que yo, con Audi y Lemos, no localizo la justificación de las creencias a priori en la comprensión de sus contenidos. Más bien, localizo la justificación de la creencia en el estado de algo que parece ser el caso.

56. Desde que escribí las observaciones en el texto, he llegado a la conclusión de que estas condiciones son necesarias pero no suficientes para la captación adecuada de un universal, aproximadamente porque, por ejemplo, el concepto de un objeto simultáneamente rojo y verde no cuenta como captación de ningún universal. No sé qué otras condiciones deben añadirse.

57. En el sentido de los filósofos, una definición debe establecer condiciones necesarias y suficientes para la aplicación de un concepto en todos los mundos posibles.

58. Pero esto no requiere la existencia de universales disyuntivos. Si se disocian dos conceptos, cada concepto puede referirse a un universal sin que la disyunción se refiera a uno.

59. El "no puede" se aplica al menos a todos los mundos posibles cuyos seres inteligentes son aproximadamente como nosotros psicológicamente. Sospecho que se aplica a todos los mundos posibles, ya que rechazo la suposición humeana de que "cualquier cosa puede causar cualquier cosa". Pero creo que la afirmación modal más débil es todo lo que requiere mi explicación del conocimiento a priori.

60. En otras palabras, considero que el problema de Benacerraf se refiere a la garantía en el sentido de Plantinga (1993, p. 3) más que a la justificación.

6

Desacuerdo y error

Pretendo haber refutado, en los últimos cuatro capítulos, cuatro de las cinco teorías metaéticas: el no cognitivismo, el subjetivismo, el naturalismo y el nihilismo. Como es habitual en filosofía, las "refutaciones" son provisionales: se ha demostrado que esas teorías se enfrentan a graves problemas que justifican su rechazo mientras exista alguna alternativa plausible. El intuicionismo ético es la alternativa que queda. Pero la opinión recibida en este campo ha sido durante mucho tiempo que el intuicionismo "es una idea irremediablemente ingenua que puede descartarse en unas pocas frases. Propongo ignorar esta teoría", escribe Brandt sobre una tesis relacionada; pero la mayoría de los autores simplemente la ignoran sin previo aviso. Los libros de texto de ética contienen ahora "refutaciones" de un párrafo de la doctrina, si es que la discuten. Todo esto pone de manifiesto la estima en la que se tiene al intuicionismo hoy en día.[1] Así que, el espectador ingenuo podría concluir que los filósofos contemporáneos deben conocer algunas objeciones obvias y decisivas a la teoría. ¿Cuáles son?

La mayoría de los expertos consideran que la existencia de desacuerdos y errores morales es la principal razón para rechazar el intuicionismo. Hay tres versiones del argumento. Una es que la

existencia de desacuerdo implica que la moralidad no es objetiva. Otra es que los intuicionistas no pueden explicar de forma plausible el desacuerdo o el error. La última es que los intuicionistas no tienen ningún método para resolver los desacuerdos.

Antes de elaborar estos argumentos, revisemos el fenómeno que sus defensores tienen en mente.

6.1 La prevalencia del desacuerdo moral

Hay tres tipos principales de desacuerdo moral:

En primer lugar, los desacuerdos entre culturas. Los miembros de una tribu de Nueva Guinea solían comerse los sesos de sus parientes muertos como parte de sus ceremonias de duelo. Los sacerdotes aztecas solían realizar sacrificios humanos, arrancando los corazones aún latientes de las víctimas para dárselos de comer al dios sol. Muchas sociedades, incluida la nuestra antes del siglo pasado, han practicado habitualmente la esclavitud. Todas estas prácticas horrorizan a los observadores occidentales modernos, pero otras sociedades las han considerado aceptables o incluso obligatorias.[2]

En segundo lugar, los desacuerdos populares dentro de nuestra sociedad. Algunos creen que el aborto es un asesinato y, por tanto, uno de los peores crímenes que se pueden cometer; otros creen que no es malo en absoluto. Algunos creen que la justicia exige la pena capital para los criminales atroces; otros consideran que la pena capital es incivilizada e injusta. Algunos piensan que la discriminación positiva es necesaria para reparar injusticias pasadas; otros piensan que aumenta la injusticia.

En tercer lugar, los desacuerdos teóricos de los filósofos. Algunos

filósofos piensan que la acción correcta es siempre la que tiene las mejores consecuencias; otros dicen que el fin no justifica los medios, y que algunos tipos de acción son incorrectos aunque produzcan las mejores consecuencias. Algunos filósofos sostienen que el único bien de la vida es el placer; otros piensan que hay muchos bienes. Algunos filósofos sostienen que debemos valorar los intereses de los demás tanto como los nuestros; otros dicen que sólo nos debemos a nosotros mismos.

Estos son algunos de los muchos ejemplos que podrían darse. Pintan una imagen de fuertes desacuerdos fundamentales como la norma en cuestiones morales. Pero esto es engañoso por dos razones. En primer lugar, muchos de estos desacuerdos no son finalmente morales, sino que se deben a diferencias en las circunstancias (no morales) en las que la gente se encuentra o cree que se encuentra. Tomando uno de los ejemplos de James Rachel, en principio parece que los esquimales, al aceptar el infanticidio, tienen valores muy diferentes a los nuestros. Pero su aceptación del infanticidio se explica por las duras condiciones en las que viven, en las que la comida escasea. En estas condiciones, las familias pueden ser incapaces de mantener a todos los bebés, y el intento de hacerlo puede amenazar la supervivencia de otros. Por lo tanto, en sus circunstancias, el infanticidio tiene sentido; no tiene por qué indicar un menor respeto por la vida por parte de los esquimales, ni valores fundamentalmente diferentes a los nuestros.[3] Por poner otro ejemplo, la práctica de los sacrificios humanos tenía sus raíces en las creencias religiosas. Los pueblos primitivos pensaban que los dioses exigían sacrificios. Los aztecas creían que el sol dejaría de brillar literalmente si no se realizaban sacrificios humanos; además, creían que las almas de las víctimas de los sacrificios iban al mejor más allá... por lo que los sacerdotes en realidad les estaban haciendo un favor. Nuestra sociedad rechaza los sacrificios humanos porque rechazamos esas creencias, no porque tengamos valores fundamentalmente diferentes.

Muchas disputas morales dentro de nuestra sociedad tienen raíces no morales. Los activistas pro-vida y pro-elección pueden diferir en cuestiones como si los fetos son sensibles, si tienen alma y si Dios prohíbe el aborto. Los desacuerdos sobre la discriminación positiva pueden implicar desacuerdos sobre las causas de las desigualdades económicas entre las razas. No digo que esto explique todos los desacuerdos morales, pero reduce la prevalencia e importancia de los desacuerdos que pueden considerarse fundamentalmente morales.

La segunda razón por la que los ejemplos de desacuerdo son engañosos es que se han seleccionado deliberadamente frente a un número mucho mayor de cuestiones morales sobre las que existe un acuerdo general. Los casos del aborto, la pena capital y la acción afirmativa no se han seleccionado al azar frente a todas las cuestiones morales; se han seleccionado para que se les preste atención porque se encuentran entre los temas más notorios de controversia. El proponente del argumento del desacuerdo no nos llama la atención sobre casos paradigmáticos de asesinato -por ejemplo, matar al dueño de una tienda durante un robo a mano armada para que no se pueda identificar más tarde-. Tampoco plantea la cuestión de la maldad de la mentira indiscriminada, el robo y la ruptura de promesas. Esas cosas están más cerca del núcleo de nuestro código moral, pero no atraen la atención ni se convierten en temas de debate público. El discurso público se centra naturalmente en las áreas de desacuerdo. Esto nos da una impresión exagerada de la importancia de dicho desacuerdo.

Los desacuerdos de los filósofos también son algo exagerados. Los filósofos están de acuerdo en casos paradigmáticos de acciones ilícitas: matar a gente por diversión, torturar, casos típicos de robo. Los filósofos morales discuten entre sí en gran medida apelando a casos hipotéticos sobre los que generalmente tienen las mismas intuiciones. De hecho, si

no estuviéramos de acuerdo en la mayoría de los casos, no podríamos tener una discusión interesante, ni siquiera podríamos entender las palabras del otro, ya que si tu concepción del bien no se parece a la mía, entonces no entenderé tus palabras como si se refirieran a la bondad en absoluto, ni podré darte crédito con ningún argumento racional sobre el tema. En lo que no están de acuerdo los filósofos es en la explicación de los casos paradigmáticos. Por ejemplo, pueden diferir sobre si el asesinato es malo porque viola el derecho a la vida, o porque disminuye la cantidad total de valor en el mundo.

No se trata de negar la existencia de importantes desacuerdos sobre los valores. Sólo se trata de decir que estos desacuerdos son menos frecuentes de lo que a veces se piensa, y que tienen lugar en un contexto de acuerdo moral sustancial.

6.2 El veto del idiota

La primera versión del argumento del desacuerdo afirma que la existencia del desacuerdo sobre la moral es incompatible con el hecho de que la moral sea "objetiva".[4] No me detendré mucho en este argumento, ya que pocos filósofos serían culpables de presentarlo. La premisa del argumento es que una afirmación es "objetivamente" verdadera sólo si todo el mundo está de acuerdo o estaría de acuerdo con ella (tal vez después de escuchar los argumentos a favor).

Esta premisa introduce lo que yo llamo el "veto del idiota". El veto del idiota es el principio según el cual se puede excluir cualquier (supuesto) hecho del ámbito de la objetividad por el mero hecho de no comprenderlo. En efecto, la premisa concede a los individuos con capacidades cognitivas limitadas o disposiciones obstinadamente escépticas un poder de veto sobre cualquier supuesta verdad objetiva.

Supongamos que un profesor de física se encuentra con un alumno que se niega a aceptar la Segunda Ley de la Termodinámica (este tipo de individuos son la fuente de los eternos esfuerzos por diseñar "máquinas de movimiento perpetuo"). Según el Veto del Idiota, esto descartaría la objetividad de la Segunda Ley de la Termodinámica.

Ningún sentido interesante de "objetivo" hace que eso sea cierto. Tomemos la definición según la cual la objetividad es una cuestión de independencia de los observadores. Según esta interpretación, el Veto del Idiota afirma que para que una cosa exista independientemente de los observadores, todos deben estar de acuerdo en que existe. Esto es un non sequitur. La definición de objetividad no dice nada sobre el acuerdo.

Una segunda concepción considera que una verdad "objetiva" es aquella que ha sido o puede ser demostrada. Según esta interpretación, el Veto del Idiota sostiene que si al menos una persona no está de acuerdo con *x*, entonces *x* no se ha demostrado. Pero, según cualquier interpretación ordinaria de "prueba", es posible que una persona no acepte la conclusión de una prueba, ya sea por falta de comprensión, por error, por parcialidad o por lo que sea, y el hecho de que no crea esa conclusión no invalida la prueba. Una prueba es, a grandes rasgos, un argumento sin fallos; el mero hecho de que alguien no acepte la conclusión de un argumento no constituye un fallo del mismo.

Por supuesto, se podría estipular simplemente que "objetivo" significa que "sería aceptado por todos". Pero esto privaría al argumento de todo interés filosófico. No sólo sería un uso engañoso del término, sino que sería un uso según el cual bien podría no haber ninguna proposición que fuera "objetiva". El interés del debate sobre la objetividad moral procede de la opinión de los antirrealistas de que la ética difiere de algún modo crucial de otros campos más "objetivos", como la física y las matemáticas. El Veto del Idiota no identifica tal

diferencia. Implica que la Segunda Ley de la Termodinámica no es objetiva, como hemos visto. Implica que la Teoría de la Evolución no es objetivamente cierta, puesto que los creacionistas la rechazan. E implica que la existencia misma del mundo material no es objetiva, puesto que el obispo Berkeley la rechaza.

Para que esta versión del argumento del desacuerdo resulte plausible, debemos matizar su premisa. Quizá la premisa debería ser que si una persona plenamente racional e informada no está de acuerdo con una afirmación, ésta carece de objetividad.[5] Pero las personas racionales e informadas cometen errores. Una persona que comete un error de cálculo, o que malinterpreta o no entiende un concepto, no por ello es irracional, está mal informada o desinformada. Es de suponer que el hecho de que alguien cometa un error sobre *x* no refuta la objetividad de *x* en ningún sentido interesante, por lo que hay que matizar de nuevo la premisa. También podemos añadir una matización relativa a la parcialidad, ya que no estar de acuerdo con *x* debido a la parcialidad no debería impugnar la objetividad de *x*. Nuestra nueva versión de la premisa (que ya no es el Veto del Idiota) es la siguiente: Si una persona totalmente racional, informada e imparcial, que entiende correcta y completamente *x* y no comete errores, no está de acuerdo con *x*, entonces *x* no es objetivamente verdadera.

Quizá la premisa de esta última versión sea correcta. Pero ahora el problema es demostrar que existen desacuerdos morales del tipo pertinente. Aunque podemos observar fácilmente desacuerdos sobre cuestiones morales, pocos seres humanos, si es que hay alguno, son completamente racionales, informados o imparciales. ¿Y cómo establecer que ambas partes de un desacuerdo comprenden correcta y plenamente la cuestión y no han cometido errores? Hay muchas formas en las que los seres humanos pueden cometer errores, la mayoría de las

cuales son comunes en cuestiones morales, como veremos en la siguiente sección.

6.3 ¿Pueden los intuicionistas explicar el desacuerdo?

La segunda versión del Argumento para el Desacuerdo afirma que la evidencia del desacuerdo moral favorece fuertemente al antirrealismo sobre el intuicionismo, porque los antirrealistas son más capaces de explicar el desacuerdo que los intuicionistas. Los antirrealistas entienden que los desacuerdos interculturales se deben a las diferentes formas de vida, junto con la ausencia de hechos que determinen la forma de vida "correcta". Entienden que los desacuerdos interpersonales se deben a las diferentes preferencias personales, junto con la ausencia de hechos que determinen las preferencias "correctas". El intuicionista, por el contrario, debe decir que la gente está sujeta a un error generalizado, pero no tiene ninguna explicación de por qué esto sería así. En resumen:

1. Hay un desacuerdo generalizado sobre los valores.
2. Si el intuicionismo fuera cierto, no habría (probablemente) un desacuerdo generalizado sobre los valores.
3. Por lo tanto (probablemente), el intuicionismo es falso.[6]

6.3.1 La caricatura

¿Cómo se supone que el intuicionismo excluye la posibilidad de un error generalizado? He aquí una respuesta popular; la llamo "la caricatura":

Caricatura Los intuicionistas caricaturescos sostienen: (i) Todas las cuestiones morales pueden resolverse apelando directamente a la intuición. No hay lugar para el razonamiento o cualquier otro proceso de investigación en

> la filosofía moral. (ii) La intuición es infalible. Y (iii) todo el mundo sabe estas cosas. De ello se deduce que todo el mundo puede obtener inmediatamente una respuesta a cualquier cuestión moral simplemente consultando su intuición. Estas respuestas deben coincidir todas, ya que la intuición es infalible. Y como todo el mundo conoce estos hechos, nadie podrá formarse una creencia moral por ningún otro medio. Entonces, ¿todo el mundo estará de acuerdo.[7]

Si la historia de la filosofía ha visto una lectura errónea menos caritativa de una posición filosófica importante, no sé cuál es. No conozco a nadie que haya sostenido los puntos de vista que se atribuyen al intuicionismo. Los intuicionistas han negado repetidamente la pretensión de infalibilidad en particular.[8] Nowell-Smith apoya la atribución con la siguiente cita del intuicionista del siglo XVIII Samuel Clarke:

> *Estas cosas son tan notoriamente claras y evidentes, que nada, salvo la extrema estupidez de la mente, la corrupción de los modales o la perversidad del espíritu, puede hacer que alguien tenga la menor duda al respecto*.[9]

Nowell-Smith produce esta cita mientras discute la supuesta incapacidad de los intuicionistas para aceptar el desacuerdo moral; se olvida de informar a sus lectores a qué se refiere "estas cosas" en el pasaje, dejándoles quizás inferir que se refiere a todas las verdades morales. En realidad, Clarke se refiere a una serie específica de principios morales que ha enumerado, principios como los siguientes:

> *[En] el trato y la conversación de los hombres entre sí, es innegable que es más conveniente... que todos los hombres se esfuercen por promover el bien universal y el bienestar de todos, que los hombres estén continuamente tramando la ruina y la destrucción de todos.*[10]

Aunque hay asuntos morales importantes sobre los que la gente no está de acuerdo, el principio anterior no está entre ellos. Los únicos ejemplos de Clarke que son controvertidos son los que se refieren a las obligaciones de y hacia Dios, donde la fuente de controversia es metafísica, en lugar de ética. Por tanto, el pasaje citado no justifica atribuir a Clarke la opinión de que somos infalibles en todas las cuestiones morales, por no hablar de otros intuicionistas.

En cualquier caso, si alguien ha sostenido los puntos de vista absurdamente fuertes y simplistas descritos en la caricatura, yo no lo hago.

6.3.2 La prevalencia del desacuerdo no moral

Quizás he sido poco amable con el argumento del desacuerdo. Una versión más sutil sostiene, no que el intuicionismo descarte cualquier desacuerdo moral, sino que el intuicionismo hace improbables desacuerdos del tipo y número que encontramos en la filosofía moral. Los desacuerdos en filosofía moral están muy extendidos, a menudo son aparentemente irresolubles y a veces se refieren a cuestiones fundamentales del tipo de las que esperaríamos que la intuición respondiera. Es este patrón específico de desacuerdo el que los intuicionistas supuestamente no pueden explicar. Aunque rechazan la infalibilidad, presumiblemente los intuicionistas creen al menos que la intuición es generalmente fiable a la hora de producir creencias correctas; de lo contrario, difícilmente podría explicar el conocimiento moral. Pero el patrón observado de desacuerdo moral impugna la fiabilidad de la intuición moral y el razonamiento moral, ya que si tuviéramos disponible un método fiable para identificar las verdades morales, debería haber relativamente pocos desacuerdos, y los desacuerdos deberían resolverse con frecuencia mediante la aplicación de ese método.[11]

No estoy de acuerdo. No creo que la disponibilidad de conocimiento moral sea incompatible con desacuerdos morales frecuentes y persistentes. En esta subsección se expone la primera parte de mis argumentos al respecto: la prevalencia de desacuerdos no morales con las mismas características problemáticas que los desacuerdos morales. Consideremos algunos ejemplos:

1. ¿Quién disparó a FK? Cuarenta años después del asesinato, muchos siguen convencidos de que hubo una conspiración, en la que posiblemente estén implicados cubanos, la mafia o la CIA. Otros están igualmente convencidos de que el responsable fue un pistolero solitario, Lee Harvey Oswald.

2. ¿Cuál es el mejor equipo de fútbol de Estados Unidos?

3. ¿Cómo debe interpretarse la mecánica cuántica? Einstein y Bohr debatieron sin problemas si el formalismo de la mecánica cuántica era completo. A día de hoy, siguen existiendo múltiples interpretaciones de la mecánica cuántica, cada una con sus propios defensores.

4. ¿Cuáles son las prácticas reales de otras culturas? Durante décadas, los antropólogos consideraron que había muchas culturas en las que el canibalismo era común. En 1979, William Arens argumentó que esto era un mito, que de hecho ninguna cultura había practicado el canibalismo de forma generalizada. Los antropólogos discreparon fuertemente en sus valoraciones del trabajo de Arens. Del mismo modo, surgió un acalorado debate cuando en 1983 Derek Freeman pretendió demostrar que el famoso relato de Margaret Mead sobre la vida en Samoa era totalmente inexacto. Ambas controversias despertaron fuertes emociones y continúan hasta el día de hoy.[12]

5. ¿Cuáles son los efectos económicos de los programas sociales del gobierno? Algunos piensan que han sido esenciales para mejorar la

gran pobreza y el sufrimiento que, de otro modo, sufrirían las clases bajas en un país capitalista. Otros piensan que esos mismos programas en realidad perjudican a los pobres.[13]

6. ¿Qué religión, si es que hay alguna, es la correcta? Hemos tenido milenios para pensar en ello, y ni siquiera nos hemos puesto de acuerdo en si existe un Dios, y mucho menos en lo que quiere.

7. ¿Existen fenómenos paranormales? Sospecho que la mayoría de los lectores responderán que no. Pero una cuarta parte de los estadounidenses cree en la astrología, la mitad cree en la percepción extrasensorial y una cuarta parte cree haber tenido una experiencia telepática.[14]

8. ¿Qué causa la enfermedad? Las mismas culturas que tienen creencias morales muy distintas de las nuestras también aceptan respuestas muy diferentes a esta pregunta. Muchos piensan que las enfermedades pueden ser causadas mágicamente por la mala voluntad de otros. La sociedad occidental aceptó durante siglos que las enfermedades estaban causadas por desequilibrios de los cuatro humores del cuerpo.

Se podrían dar muchos más ejemplos. Los desacuerdos de los ejemplos 1-7 tienen las siguientes características:

i) Son generalizados entre las personas que tienen una opinión sobre los temas; no son sólo unas pocas personas las que van en contra de un consenso.

ii) Son aparentemente irresolubles. Cualquiera que piense que puede resolver quién disparó a JFK "simplemente entrando y mirando los hechos" está invitado a intentarlo. Es posible que descubra "la respuesta" (según su criterio), pero sólo los no iniciados imaginarán que tienen una oportunidad de convencer a la(s) otra(s) parte(s).

iii) Se refieren a hechos objetivos, no valorativos. De hecho, cuestiones

como quién disparó a JFK o si una determinada cultura practicaba el canibalismo son paradigmas de preguntas objetivas.

En el caso del ejemplo 8, hubo una respuesta consensuada en nuestra sociedad durante varios siglos, pero esta respuesta era completamente errónea.

He enumerado una serie de ejemplos para ilustrar el alcance del desacuerdo humano. La gente no está de acuerdo sobre los acontecimientos históricos, sobre cuestiones aparentemente sencillas de hechos empíricos, sobre cosas supuestamente establecidas por la ciencia, sobre la interpretación de teorías aceptadas y sobre muchas cuestiones filosóficas, entre otras cosas. Incluso cuando estamos de acuerdo, eso no garantiza que tengamos razón, como muestra el ejemplo 8. Es cierto que he dado una visión parcial de nuestra vida intelectual: he seleccionado intencionadamente casos notables de desacuerdos fuertes y persistentes. Pero el antirrealista ha hecho lo mismo en el ámbito moral al citar ejemplos como el aborto y la pena capital.

Los desacuerdos persistentes en el ámbito no moral son comunes. Sin embargo, prácticamente nadie considera que esto refute la posibilidad del conocimiento en campos como la historia, la antropología y la economía. Sospecho que pocos sacarían la conclusión, incluso ligeramente escéptica, de que no podemos adquirir conocimientos sobre las cuestiones específicas en las que no estamos de acuerdo (aunque todavía no lo hayamos hecho). Entonces, ¿cómo se supone que el desacuerdo moral refuta la posibilidad del conocimiento moral?

Quizá haya algo especialmente problemático en los desacuerdos en áreas en las que el conocimiento depende de intuiciones, a diferencia de áreas en las que el conocimiento depende de otros medios, como la observación y el razonamiento científico. Si lo hay, no tengo ni idea de lo que es. Los intuicionistas no tienen por qué sostener ni sostienen que

el conocimiento basado en intuiciones sea superior al conocimiento empírico; no tienen por qué considerar que el razonamiento basado en intuiciones morales sea más fiable que la ciencia, por ejemplo. Quienes se oponen al intuicionismo no hacen más que aplicar un doble rasero epistemológico: Si el razonamiento científico y otros razonamientos empíricos no llegan a un acuerdo en muchos casos, pero esto no los invalida como fuente de conocimiento, ¿por qué el hecho de que la intuición tampoco llegue a un acuerdo debería invalidarla como fuente de conocimiento?

Tal vez el problema sea que en el caso de la filosofía moral tenemos muchas disputas dentro de la misma materia, mientras que las disputas no morales son casos aislados distribuidos en muchas materias diferentes. Pero esto no es así. Hay numerosas disputas sobre el tema de la religión, incluso sus afirmaciones no morales; todas las ramas de la filosofía; los deportes, en particular los méritos de varios jugadores y equipos y lo que la gente "debería" hacer (en un sentido no moral) en varios puntos de un juego; las habilidades relativas de individuos específicos que uno puede conocer (por ejemplo, ¿es Sally más inteligente que yo?); y todo lo que tiene que ver con la política, de nuevo incluyendo cuestiones no morales. Sin embargo, en ninguno de estos ámbitos no existen hechos objetivos, ni carecemos de facultades cognitivas capaces de aprehender hechos de ese tipo.

6.3.3 Una colección de errores

¿Qué hay de sorprendente en el desacuerdo, de tal manera que requiera una explicación especial? Las creencias, pensamos, apuntan a la verdad. Si las personas alcanzan la verdad, entonces estarán de acuerdo entre sí. Así que el desacuerdo generalizado implica un error generalizado, aunque aparentemente intentemos evitar el error. ¿Por qué no dejamos

de utilizar los métodos poco fiables de formación de creencias que nos llevan a los desacuerdos? ¿Por qué no nos sentamos, con nuestros libros de texto de lógica y pensamiento crítico en la mano, y averiguamos quiénes son los que presentan los malos argumentos?

Las preguntas anteriores se basan en suposiciones ingenuas. Aunque las creencias en cierto sentido apuntan a la verdad, las creencias de las personas pueden gratificar muchas otras pasiones además del anhelo racional de la verdad. Los cánones de la lógica y el pensamiento crítico tampoco determinan algorítmicamente qué es y qué no es un mal argumento. Las personas rara vez reúnen las condiciones ideales para razonar, e incluso las mejores condiciones cognitivas no hacen infalibles a los seres humanos. Las causas de nuestras creencias pueden ser muy complejas y ocultársenos en gran medida. A menudo no podemos articular algunas de las razones de nuestras creencias. El intuicionista no puede crear un algoritmo para calcular la teoría moral correcta, como tampoco nosotros podemos crear un algoritmo general para calcular teorías empíricas correctas.

Los críticos han afirmado que el intuicionista no dispone de otra explicación del error moral que no sea la "ceguera", la estupidez o la mendacidad deliberada.[15] Para evaluar mejor esta acusación, revisemos algunas de las fuentes de error en general:

1. *Prejuicios*: Las personas que tienen prejuicios están movidas por emociones y deseos, además del deseo de tener creencias verdaderas y evitar las falsas, para ver algunas afirmaciones de forma favorable y otras de forma desfavorable. En otras palabras, hay cosas que queremos creer y otras que no queremos creer. A menudo, queremos creer lo que nos interesa creer o convencer a los demás; lo que nos gustaría que fuera cierto; lo que le interesa creer al grupo social con el que nos identificamos; y lo que es coherente con la imagen que

queremos mantener de nosotros mismos. Estos son sólo algunos de los posibles sesgos que podemos sufrir.[16]

2. *Error de cálculo*: Mientras realizo un cálculo, me olvido de llevar un "1", lo que da lugar a una respuesta incorrecta. Si nadie me lo indica, es posible que nunca me dé cuenta de mi error. Puedo entender perfectamente los principios matemáticos implicados en el problema, ser racional, imparcial e inteligente, y aun así cometer un error.

3. *Confusión*: Se trata de no distinguir dos cosas superficialmente similares. Por ejemplo, el estudiante que piensa que como no podemos hablar de nada sin usar el lenguaje, no podemos hablar del mundo independiente del lenguaje, está cometiendo una confusión.

4. *Malentendidos y falta de comprensión*: Los conceptos y las teorías pueden ser malinterpretados. Un estudiante que piense que la teoría del caos da cabida al libre albedrío sufre una falta de comprensión del concepto de caos y quizás del libre albedrío. El filósofo que considera que el intuicionismo es absurdo porque no da cabida al debate moral sufre de una incomprensión de la teoría.

5. *Supervisión*: La gente puede no darse cuenta de consideraciones relevantes, incluyendo algunas de las implicaciones de sus afirmaciones. Alguien me pregunta si creo que dos objetos sólidos pueden superponerse en el espacio. Yo respondo que no, basándome en mi impresión inmediata, y pasando por alto ejemplos como mi mano y mi cuerpo, o una estatua y la masa de arcilla que la compone. Por otra parte, las personas pueden ejercer una atención selectiva, es decir, pueden atender sólo a algunas de las características relevantes de una situación, ya sea por prejuicios o simplemente porque la situación es demasiado compleja para asimilarlo todo a la vez.

6. *Juicios precipitados*: Las personas suelen emitir un juicio basado en una apariencia inicial, cuando una reflexión más larga y cuidadosa habría dado lugar a un juicio diferente.

7. *Información falsa o incompleta*: A veces las personas no conocen todos los hechos relevantes para emitir un juicio, y a veces hacen suposiciones erróneas. Por ejemplo, dos personas que discuten sobre el control de armas pueden haber recibido información radicalmente diferente, una consultando la literatura de la NRA, y la otra consultando la literatura de Handgun Control, Inc. Parte de su información puede ser falsa, mientras que ninguna de ellas tiene información completa.

8. *Afirmaciones no articuladas*: Imaginemos que dos personas discuten sobre si el aborto debe ser legal. Una de ellas supone que está discutiendo sobre si el aborto debería estar permitido alguna vez, mientras que la otra supone que está discutiendo sobre si debería estar permitido siempre. O tal vez ninguno de los dos sabe realmente cuál de estas cuestiones, si es que alguna, pretenden discutir.

9. *La terquedad*: La gente puede simplemente negarse a reconsiderar sus viejas creencias, incluso cuando aparecen nuevas pruebas relevantes.

10. *Falacias*: Son ilusiones de la razón, formas de razonamiento que pueden parecer superficialmente convincentes pero que en realidad no lo son; por ejemplo, el equívoco, la afirmación del consecuente, el argumento ad hominem. Especialmente importante es la confianza errónea en la autoridad: la gente puede creer una cosa porque otros -padres, líderes religiosos, periodistas u otros miembros de su sociedad- lo han afirmado. Pero esos otros pueden no saber de qué están hablando.

11. *Olvido*: Las personas pueden olvidar o simplemente no darse cuenta de sus propias razones para mantener una creencia, y

pueden olvidar otras pruebas relevantes. En este caso, están en mala posición para reconocer cuando la creencia ha sido socavada.

12. *Dificultad intrínseca de las cuestiones*: Algunas cuestiones, como las que son especialmente abstractas o complejas, son más difíciles de comprender con claridad y de emitir juicios que otras.

13. *Pruebas inarticuladas*: A veces tenemos motivos para creer que no podemos articular. Supongamos que una amiga ha sido acusada de malversar dinero de su empresa. Como la conozco bien, sé que no haría algo así. Pero no podría demostrarlo a un tercero. Este tipo de pruebas inarticuladas pueden provocar desacuerdos.

14. *Defectos mentales*: Por último, y quizá el menos importante, algunas personas carecen de las capacidades cognitivas que tienen la mayoría de los demás, o tienen capacidades cognitivas cuyo funcionamiento está gravemente distorsionado, como en el caso de los esquizofrénicos, los autistas y los psicópatas. Aunque un esquizofrénico paranoico pueda parecer un razonador competente, no tendrás suerte razonando para sacarlo de sus delirios.

Estas fuentes de error no son independientes unas de otras; algunas se solapan y otras influyen en otras.

No puedo afirmar que mi colección de errores sea completa; tal vez haya otras fuentes de error que haya pasado por alto. Dada la multitud de formas en que los seres humanos pueden equivocarse, ¿por qué los antirrealistas piensan que el desacuerdo moral debe explicarse por una escasez de hechos objetivos? Todas las fuentes de error y desacuerdo mencionadas anteriormente operan en asuntos no morales cuya objetividad es indiscutible; no hay ninguna razón obvia por la que no deban operar también en asuntos morales.

6.3.4 El desacuerdo es predecible

Los defensores del argumento del desacuerdo, demasiado ansiosos en sus rechazos del intuicionismo, han olvidado justificar la segunda premisa crucial de su argumento:

2. Si el intuicionismo fuera cierto, no habría un desacuerdo generalizado sobre los valores.

Como hemos visto, hay muchas maneras en que los seres humanos pueden equivocarse, y hay muchos desacuerdos persistentes sobre hechos objetivos, en principio conocibles. A la luz de esto, no tenemos motivos obvios para predecir que no habrá disputas generalizadas sobre los valores.

Pero ahora quiero decir algo más fuerte: creo que tenemos motivos obvios para predecir que habría disputas generalizadas sobre los valores, incluso si el intuicionismo fuera cierto. Para ello es necesario observar los tipos de casos en los que el desacuerdo y el error son comunes, en general. Hay al menos cuatro tipos de casos a considerar.

Primero: El desacuerdo y el error son comunes cuando hay sesgos fuertes y frecuentes. El sesgo es especialmente importante, ya que aumenta la frecuencia y la gravedad de la mayoría de los demás tipos de errores, como los descuidos, los malentendidos, los olvidos, etcétera. Probablemente desempeñe un papel importante en los desacuerdos enumerados en la sección 6.3.2. He aquí algunas fuentes comunes de sesgo, todas las cuales pueden afectar a cuestiones morales:

El interés propio: A mí me interesa, por ejemplo, creer que las cosas que quiero hacer son moralmente aceptables, y que las cosas que quiero que hagan los demás son obligatorias. Incluso me interesa convencer a los demás de estas cosas.

Deseo: Me gustaría que fuera cierto que el mundo es justo; que las cosas son buenas como son; que yo y mi grupo social somos

especialmente dignos de elogio; que nuestras prácticas están justificadas.

Intereses de grupo: A mi grupo social le interesa creer que somos superiores a otros grupos, que las cosas que queremos hacer están justificadas, etc.

Autoinclusión: A veces, la imagen que queremos construir de nosotros mismos afecta a las creencias que adoptamos. Si quiero presentarme como compasivo e indulgente, puede que quiera rechazar la pena capital; pero si prefiero presentarme como duro e inflexible, puede que quiera aceptar la pena capital.

Segundo: El desacuerdo y el error son comunes en áreas donde la gente se atiene a las tradiciones de sus culturas. Las tradiciones culturales son una fuente de información notoriamente poco fiable. Las culturas con creencias morales sorprendentes, que se supone que nos hacen dudar de la realidad de los hechos morales, también tienen una serie de creencias no morales sorprendentes -sobre cosas como las causas de las enfermedades, la estructura del cosmos, la historia humana y el funcionamiento del mundo físico- que sabemos, científicamente, que son falsas.

Tercero: el desacuerdo y el error son comunes en áreas en las que la gente se remite a la religión, otra fuente de información poco fiable. En la sección 3.4, mencioné algunas creencias morales falsas apoyadas por la Biblia. He aquí algunas falsas creencias fácticas también aparentemente apoyadas por la Biblia:

Los conejos rumian (*Deuteronomio* 14:6-7)

Algunos insectos tienen cuatro patas (*Levítico* 11:20-22)

La Tierra es plana (*Mateo* 4:8)

$\pi = 3$. (*1 Reyes* 7:23)

La Tierra es inmóvil y descansa sobre un fundamento. (*Salmos* 104:5)

La Tierra tiene unos 6000 años de antigüedad. (Supuestamente apoyado por varias partes de la Biblia.[17])

Cuarto: El desacuerdo es común en todas las áreas de la filosofía. Durante siglos, los filósofos han estado en desacuerdo sobre cosas como Dios, el libre albedrío, los universales, el conocimiento a priori y la naturaleza de la mente, por nombrar algunas. El nivel de desacuerdo en la ética no es mayor que el encontrado en la metafísica, la epistemología, la filosofía de la mente o la filosofía del lenguaje.[18]

En los cuatro aspectos, deberíamos esperar que las cuestiones morales sean especialmente propensas al desacuerdo. Mi respuesta a la principal premisa del argumento del desacuerdo, por tanto, es la siguiente: incluso si hubiera verdades morales objetivas y tuviéramos formas disponibles de conocerlas, todavía cabría esperar desacuerdos generalizados, fuertes y persistentes. No se requiere ningún supuesto antirrealista para predecir el fenómeno.

6.4 ¿Pueden los intuicionistas resolver los desacuerdos?

La versión final del Argumento del Desacuerdo es algo así:

1. El intuicionismo no proporciona ninguna forma de resolver los desacuerdos morales.
2. Si una teoría metaética no proporciona ninguna forma de resolver los desacuerdos morales, entonces es falsa.
3. Por lo tanto, el intuicionismo es falso.

Normalmente, sólo la premisa (1) se enuncia explícitamente. La premisa

(2) también es necesaria, pero su justificación sigue siendo oscura.

6.4.1 Desacuerdos hipotéticos

¿Cómo debemos interpretar (1) ¿Qué desacuerdos se le reprochan al intuicionista por su incapacidad para resolverlos? Para algunos críticos, la queja parece ser que el intuicionista no tiene forma de resolver ciertos desacuerdos hipotéticos. Así, Nowell-Smith estipula un caso en el que alguien no está de acuerdo con nosotros sobre una cuestión moral y no acepta ninguna premisa que podamos utilizar para apoyar nuestro punto de vista. En tal caso, dice, no hay forma de resolver la disputa mediante el razonamiento.[19] Aunque no cabe duda de que tiene razón, esto no significa que el intuicionismo sea defectuoso. Nadie puede proporcionar una forma de resolver racionalmente cualquier tipo de disputa si estipulamos que cada parte se niega a aceptar cualquier premisa que la otra parte pueda utilizar. ¿Cómo puede suponer esto un problema?

6.4.2 Cuestiones morales controvertidas

Quizá la queja sea que hay desacuerdos reales que los intuicionistas no pueden resolver, como las disputas sobre el aborto, la pena capital y la discriminación positiva. Si bien esto es plausible, es al menos igual de plausible pensar que nadie más puede ofrecer tampoco una forma de resolver tales disputas. ¿Por qué deberíamos pensar que una teoría metaética sólo puede ser verdadera si proporciona una forma de resolver todas las disputas? Quizá se supone que esta es una condición de adecuación sólo para las teorías realistas, pero incluso en ese caso sigue sin estar motivada. Supongamos que acabo de exponer una teoría filosófica según la cual existen hechos objetivos sobre los objetos físicos, y todo nuestro conocimiento de tales hechos deriva de la percepción. ¿Sería justa la siguiente objeción: "No has proporcionado una manera de

determinar quién disparó a JFK. Por lo tanto, su teoría es errónea"?

6.4.3 Controversias morales fundamentales

Quizás la queja es que algunos desacuerdos sobre afirmaciones morales fundacionales -afirmaciones morales que no dependen de otras creencias- no pueden resolverse. La disputa sobre JFK se refiere a creencias que dependen del razonamiento, no a creencias fundacionales, por lo que no es análoga.

Si este es el argumento, debemos señalar en primer lugar que las disputas morales fundacionales son relativamente raras; las disputas sobre el aborto y la discriminación positiva no son casos, ya que nuestras creencias sobre esas cuestiones dependen de otras creencias. Por ejemplo, la mayoría de los que piensan que el aborto está mal lo hacen porque creen que los fetos son personas, que en general está mal matar a las personas y que el aborto mata a los fetos. Además, la literatura está repleta de ejemplos de argumentos morales racionales relativos al aborto y a otros asuntos notorios de controversia, lo que demuestra que no es necesario que nos echemos las manos a la cabeza cuando nos enfrentamos a tales disputas.[20] Aunque Ayer afirma que los desacuerdos morales fundamentales son "notorios", no proporciona ningún ejemplo. Los ejemplos más plausibles son cosas sobre las que, en su mayoría, sólo discuten los filósofos, como si debemos ponderar los intereses de otras personas por igual que los nuestros, o si la acción con mejores consecuencias es siempre la correcta.

En segundo lugar, mientras nuestro fundacionalismo tenga una forma moderada, no es imposible razonar sobre principios morales intuitivos. Podemos mantener incluso principios morales intuitivos abiertos a la revisión a la luz de nuevas pruebas. Los ejemplos citados en el párrafo anterior son principios que intuimos, pero los filósofos morales no han

renunciado a razonar sobre ellos. Volveré sobre este punto en la siguiente sección.[21]

En tercer lugar, aunque la incapacidad de resolver disputas es un grave problema práctico, simplemente no está claro cómo se supone que constituye una prueba de que la teoría del intuicionismo es falsa.

En cuarto lugar, no está claro por qué el razonamiento no está sujeto a las mismas exigencias que otros medios de cognición. Hay muchas disputas aparentemente irresolubles sobre creencias que dependen del razonamiento, pero eso no nos convence de que el razonamiento no sea un medio legítimo de cognición. Entonces, ¿por qué las disputas aparentemente irresolubles sobre creencias que dependen de la intuición nos convencerían de que la intuición no es un medio legítimo de cognición? Quizá el argumento presupone que tenemos formas sencillas de determinar, en todos los casos, si un razonamiento es convincente. Pero, por supuesto, eso es falso. No existe un conjunto de reglas generalmente aceptadas, exhaustivas y útiles sobre qué cuenta como prueba de qué. Eso es lo que ilustran ejemplos como la disputa JFK: nadie puede establecer, de forma que todas las partes inteligentes estén de acuerdo, si las pruebas disponibles apoyan una teoría de la conspiración.

Por último, imaginemos una disputa sobre una afirmación fundacional no intuitiva. En un juicio, dos testigos dan versiones distintas de los hechos y ambos afirman haber visto lo que cuentan. Este tipo de caso no es inusual. Ambos testigos pueden dar todas las muestras conocidas de creer sinceramente lo que dicen.[22] ¿Es tarea de una teoría filosófica de la percepción proporcionar una forma de resolver esa disputa? Si no puedo hacerlo, ¿queda refutada mi teoría de la percepción, es decir, mi afirmación de que el mundo físico es objetivo o mi afirmación de que la percepción es una fuente legítima de conocimiento? Si no es así, ¿cómo

podría refutar la existencia de valores objetivos o la afirmación de que la intuición ética es un medio legítimo de conocimiento el hecho de no poder resolver algunas controversias morales?

6.4.4 Controversias en general

Algunos acusan a los intuicionistas de no tener forma de resolver ninguna disputa, o al menos ninguna disputa sobre intuiciones.[23] Ante el desacuerdo, los intuicionistas deben sacar inmediatamente los garrotes y empezar a golpear a la gente en la cabeza.[24] A diferencia de las otras, esta acusación identificaría un grave defecto en el intuicionismo, ya que de hecho algunas disputas son resolubles. Pero la acusación es falsa.

Ante un desacuerdo, un intuicionista puede, entre otras cosas, intentar determinar si una de las causas comunes de error discutidas en la sección 6.3.3 es operativa, y remediar la situación si lo es. Puede investigar cuestiones como si una de las partes es parcial, si hay malentendidos o confusiones sobre alguno de los conceptos o puntos de vista implicados en el debate, si una de las partes se ha precipitado o ha cometido un descuido, si alguna de las partes tiene información falsa o incompleta, etc. Cualquiera de estos problemas podría afectar negativamente a los juicios intuitivos. Además, algunas intuiciones pueden resultar depender (causalmente) de otras creencias de manera que uno no era consciente inicialmente. A veces, esto se puede poner de manifiesto al considerar otros escenarios hipotéticos ("¿Seguiría diciendo que *x* es incorrecto si *p* fuera el caso?"), y entonces se pueden examinar estas creencias adicionales.

El problema de la parcialidad es el más preocupante. Aunque la parcialidad está omnipresente en la vida intelectual, las acusaciones de parcialidad rara vez hacen avanzar el debate. Lo más probable es que cada parte busque argumentos para demostrar que la otra es más

tendenciosa, y es probable que estos argumentos se evalúen con al menos tanta parcialidad como la cuestión moral original. Por esta razón, puede que haya poco que podamos hacer contra los prejuicios de los demás. Pero podemos intentar contrarrestar nuestros propios prejuicios. La mera identificación de un prejuicio es un gran paso para contrarrestarlo. Al abordar una cuestión moral, debemos pensar si estamos sujetos a los prejuicios antes mencionados. En este sentido, deberíamos preguntarnos si tenemos las creencias morales que tenemos debido a nuestra cultura o religión, y si es así, deberíamos sospechar de esas creencias. El hecho de que una creencia esté respaldada por nuestra cultura o religión no es prueba de que sea falsa, pero sí es prueba de que nuestra creencia no es fiable, de que probablemente la mantendríamos tanto si fuera cierta como si no.

Un ejemplo de Peter Singer ilustra este punto. Después de defender la permisibilidad del aborto, Singer se enfrenta a la acusación de que sus argumentos no pueden distinguir entre fetos y recién nacidos y, por tanto, implican que el infanticidio también es permisible en muchos casos (aproximadamente, porque los recién nacidos no tienen capacidades mentales más sofisticadas que los fetos). Responde, en parte, informando al lector de que muchas sociedades han considerado el infanticidio aceptable e incluso obligatorio en ciertos casos. Platón, Aristóteles y Séneca recomendaban matar a los bebés deformes. ¿Qué quiere decir Singer con esto? ¿Apela a la autoridad para defender el infanticidio? No; lo que quiere decir es que nuestro horror ante la idea de matar a los niños es probablemente más un producto de nuestra cultura particular (en particular, piensa él, del cristianismo) que de los hechos objetivos sobre los niños. También sugiere que una buena parte de la causa de nuestra creencia es la ternura de los bebés, que es moralmente irrelevante.[25]

No intento resolver las disputas sobre el aborto y el infanticidio. Mi

objetivo es simplemente ilustrar la importancia de mirar más allá de la propia cultura. Singer tiene razón al insistir en que la naturaleza culturalmente específica de las creencias sobre el infanticidio disminuye su autoridad.

Reducir los prejuicios es sólo una parte de una estrategia razonable para mejorar el desacuerdo moral. Los argumentos morales también deben desempeñar un papel importante en el esfuerzo por resolver las disputas, incluso las disputas sobre principios intuitivos. Los filósofos pueden razonar entre sí sobre un principio moral, por ejemplo, considerando su coherencia con otros principios morales; buscando consecuencias del principio que hayan pasado desapercibidas; considerando cómo se aplica a situaciones específicas sobre las que tenemos intuiciones; estableciendo analogías entre el principio y otros principios que aceptamos o rechazamos; y buscando restricciones generales plausibles en las teorías morales (por ejemplo: Una teoría moral debe hacer comprensible cómo la gente puede estar motivada por el bien" o, "Una teoría moral debe implicar que cuando *A* es mejor que la banda *B* es mejor que *C*, *A* es mejor que *C*").[26] Los filósofos no han renunciado ni mucho menos a resolver las cuestiones morales, aunque reconocemos que el proyecto es a largo plazo.[27] Los intuicionistas, naturalmente, adoptarán todos o casi todos los métodos actuales de la ética normativa. A la luz de esto, la mayoría de los eticistas no pueden rechazar el intuicionismo sobre la base de que no permite un enfoque razonable de las cuestiones morales controvertidas.

¿Por qué se piensa que el intuicionismo impide razonar sobre intuiciones? Podría pensarse que los fundacionalistas no pueden permitir el razonamiento sobre creencias fundacionales. Pero esto es un malentendido. Una creencia fundacional se define como aquella que no necesita una razón para estar justificada hasta cierto punto; no se define

como una creencia que no puede tener una razón. Además, la mayoría de los fundacionalistas adoptan la noción de justificación prima facie de la que se habla en la sección 5.1 anterior. Mi principio de conservadurismo fenoménico es un ejemplo de ello: sólo sostiene que las apariencias se presumen racionalmente verdaderas, hasta que aparezcan pruebas de su falsedad o falta de fiabilidad. Los intuicionistas deberían permitir que la justificación intuitiva de una creencia moral pueda ser derrotada, bien por tensiones con otras creencias morales (propias o ajenas), bien por pruebas empíricas de su falta de fiabilidad (por ejemplo, pruebas de parcialidad). Así que los intuicionistas parecen tener al menos tan buenas perspectivas para resolver disputas como cualquiera de los filósofos morales.

Puede que se sienta un escepticismo justificado sobre las perspectivas de resolver los desacuerdos por los medios que he sugerido. Pocas personas, digan lo que digan, trabajarán concienzudamente para eliminar sus propios prejuicios. La mayoría seguirá remitiéndose irreflexivamente a su cultura y religión. Incluso los filósofos morales seguirán, en su mayoría, racionalizando sus teorías favoritas. Al menos por esta razón, el desacuerdo moral está aquí para quedarse. Esto revela un grave defecto de la naturaleza humana. Pero no revela ningún defecto en el intuicionismo ético.

6.5 El problema de la autorrefutación

El argumento del desacuerdo es de doble filo: Puede volverse contra los puntos de vista de sus defensores con la misma facilidad que contra el intuicionismo. Para empezar, obsérvese que los defensores de otras posturas metaéticas consideran que sus propios puntos de vista son objetivamente ciertos. Al afirmar sus teorías sobre la naturaleza de las

afirmaciones morales, no pretenden simplemente expresar o describir sus emociones, por ejemplo; pretenden afirmar un hecho que tenemos motivos suficientes para creer (si no hay hechos morales, es un hecho [no moral] que no hay hechos morales). Pero si el argumento del desacuerdo es sólido, entonces se refuta a sí mismo, ya que mucha gente no está de acuerdo con el argumento del desacuerdo. El argumento refutaría asimismo cualquier posición metaética, debido a la naturaleza del desacuerdo en metaética. Recordemos las tres versiones del argumento del desacuerdo:

a) *El veto de los idiotas*: Si el desacuerdo de alguien con una afirmación, en general, impide que esa cláusula sea objetivamente verdadera, entonces los desacuerdos en la metaética impiden que cualquier punto de vista metaético sea objetivamente verdadero. Del mismo modo, algunas personas rechazan el argumento del desacuerdo, por lo que no puede ser objetivamente correcto.

b) *El problema de la explicación del desacuerdo*: El desacuerdo sobre las teorías metaéticas está al menos tan extendido y es tan persistente como el que existe sobre las teorías éticas. ¿Debemos explicar esto negando la existencia de cualquier teoría metaética correcta? Del mismo modo, no hay un acuerdo general sobre si el argumento del desacuerdo es correcto; ¿sugiere esto que no hay ningún hecho sobre si es correcto?

Si los desacuerdos éticos son problemáticos para el intuicionista porque sugieren que la intuición no es fiable, entonces los desacuerdos en metaética son igualmente problemáticos para cualquier metaeticista porque sugieren que la reflexión y el razonamiento filosóficos -o cualquier método que hayan utilizado los metaeticistas- no son fiables. Muchas de las ideas y argumentos centrales en este campo han existido durante décadas, y los filósofos aún no se han puesto de acuerdo sobre su solidez.

c) *El problema de resolver el desacuerdo*: Del mismo modo, si el intuicionista carece de un método eficaz para resolver las disputas éticas, todos carecemos de un método eficaz para resolver las disputas metaéticas. Una vez más, los métodos realmente empleados por los filósofos no han tenido hasta ahora más éxito en la resolución de cuestiones metaéticas que en la resolución de cuestiones éticas.

Una de las razones es que la elección de una teoría filosófica depende, por lo general, de amplios juicios sobre la plausibilidad general realizados a la luz de numerosos argumentos no concluyentes, la relación de la teoría con otras teorías filosóficas y las diversas "ventajas y desventajas teóricas" de los puntos de vista en competencia.[28] No veo ninguna alternativa a este enfoque, pero significa que es poco probable que la mayoría de las grandes disputas filosóficas se resuelvan en un futuro previsible.

¿Cómo debemos reaccionar ante el fenómeno del desacuerdo en filosofía? Se podría llegar a la justificada conclusión de que el conocimiento filosófico es muy difícil de alcanzar y que se ha alcanzado muy poco. Por lo tanto, es posible que se adopte una actitud vacilante y modesta respecto a las propias creencias filosóficas. Lo que no debe decirse -no sólo porque sería contraproducente, sino porque es evidentemente falso- es que carecemos de las facultades cognitivas necesarias para abordar las cuestiones filosóficas. Por muy difíciles que sean y por muy propensos que seamos a equivocarnos, no estamos simplemente cerrados a las cuestiones filosóficas (en el sentido en que lo está mi gato): las entendemos y podemos argumentar racionalmente sobre ellas. Los oponentes del intuicionismo no niegan esto; si lo hicieran, no participarían en el discurso filosófico. Del mismo modo, una actitud modesta hacia las creencias morales controvertidas se apoya en el fenómeno del desacuerdo moral; lo que no se justifica es negar que las cuestiones morales tengan respuesta o que tengamos las facultades

cognitivas para abordarlas.

6.6 El desacuerdo como argumento a favor del realismo

El argumento del desacuerdo es el que suelen esgrimir los antirrealistas. Pero por las razones que hemos visto, el fenómeno del desacuerdo no proporciona ninguna prueba significativa para su punto de vista. De hecho, ahora argumentaré que el desacuerdo es más problemático para los antirrealistas que para los realistas.

6.6.1 ¿Pueden los antirrealistas explicar el desacuerdo?

Para empezar, ¿cómo pueden explicar los antirrealistas por qué discutimos con otros sobre cuestiones morales, afirmando que están equivocados, etc.? Ya hemos visto las dificultades para los subjetivistas (capítulo 3): Según el subjetivismo individualista, no puedes estar en desacuerdo con la afirmación moral de alguien a menos que pienses que está informando falsamente de sus propios sentimientos. Según el relativismo cultural, no se puede estar en desacuerdo con la afirmación moral de alguien a menos que se piense que está equivocado sobre cuáles son las costumbres de su sociedad. Es difícil ver por qué debería haber disputas generalizadas e irresolubles sobre cuáles son nuestras emociones o convenciones. Desde un punto de vista no cognitivista, ni siquiera hay proposiciones sobre las que discrepar. La gente puede tener "actitudes contradictorias",[29] pero esto no explicaría por qué, por ejemplo, uno intentaría razonar con los que se oponen a la eutanasia o decirles que están equivocados. Normalmente no se discute con la gente ni se les dice que están "equivocados" simplemente porque tienen sentimientos diferentes a los propios.

Considere tres respuestas a esto. La primera: Quizá discutimos sobre

cuestiones morales porque pensamos erróneamente que el realismo moral es cierto. Esta respuesta es válida para los nihilistas. No sirve para los subjetivistas ni para los no-cognitivistas. Nuestra creencia en los valores objetivos sólo explica que discutamos sobre cuestiones morales si consideramos que los términos morales se refieren a esas propiedades objetivas de los valores. Pero los subjetivistas o los no cognitivistas piensan que el lenguaje moral no se refiere a ninguna supuesta propiedad objetiva, sino que sólo expresa o informa de estados mentales subjetivos. Esto les deja con la extraña idea de que casi todo el mundo está fundamentalmente equivocado acerca de los significados de las palabras comunes en nuestro lenguaje y que, aunque creemos en las propiedades objetivas de valor, no tenemos ninguna palabra para ellas.

Segundo: Quizá nuestras discusiones sobre cuestiones morales no se refieran a juicios de valor fundamentales; lo que estamos discutiendo en realidad son sólo los hechos no morales de los que dependen los juicios morales. Por ejemplo, quizá quienes discrepan sobre si la eutanasia es permisible en realidad sólo discrepan sobre los efectos que tendría permitirla. Pero esta respuesta renuncia a la premisa fundamental del argumento del desacuerdo, a saber, que existen disputas morales fundamentales y generalizadas.

Tercero: Quizá discutamos sobre cuestiones morales en un esfuerzo por cambiar las actitudes no cognitivas de los demás, no por cambiar sus creencias. Pero no tratamos de cambiar las actitudes de los demás limitándonos a emularlas; cuando discutimos con otros sobre la moral, decimos y pensamos que están equivocados, y pretendemos darles razones intelectuales para que cambien de opinión. Esto podría tener sentido desde un punto de vista antirrealista sólo si pensamos que los demás están equivocados sobre algunos hechos no morales de los que dependen sus actitudes morales. En ese caso, la argumentación podría

esperar cambiar las actitudes de los demás corrigiendo sus suposiciones fácticas erróneas. Pero, de nuevo, esto requiere que los desacuerdos morales se deriven fundamentalmente de errores sobre cuestiones no morales, en contra de la suposición del argumento del desacuerdo.

Los no-cognitivistas y los subjetivistas no están, por tanto, en condiciones de insistir en la supuesta incapacidad de los intuicionistas para explicar el desacuerdo moral; además, en la medida en que estamos convencidos de que algunos desacuerdos morales son realmente fundamentales, tenemos razones para rechazar el no-cognitivismo y el subjetivismo.

La situación para los nihilistas es más compleja. Pueden explicar el desacuerdo moral apelando a nuestra extendida y equivocada creencia en el realismo, pero sólo si pueden explicar por qué la gente cometería tal error. A primera vista, la explicación más sencilla de por qué casi todo el mundo cree que el realismo moral es cierto, es que es cierto. Los nihilistas han propuesto explicaciones alternativas,[30] pero nos llevaría demasiado lejos discutirlas; por ahora, entonces, permitamos que los nihilistas puedan explicar el desacuerdo moral.

6.6.2 ¿Pueden los antirrealistas resolver los desacuerdos?

Lo máximo que pueden afirmar los antirrealistas es que los intuicionistas no tienen forma de resolver los desacuerdos morales fundamentales, es decir, los desacuerdos morales que no dependen de errores de hecho o de razonamiento. Supuestamente, los intuicionistas sólo pueden exhortar a sus interlocutores a que se concentren más en la cuestión moral, y tal vez golpearlos en la cabeza. Pero, ¿qué puede ofrecer el antirrealista? Para los antirrealistas, no hay hechos valorativos objetivos; así que cuando se trata de disputas sobre valores fundamentales, no hay nada cognitivo que podamos hacer; todo lo que podemos hacer es, quizás, engatusar,

amenazar o acordar el desacuerdo. (Los nihilistas se limitan a decir que todo el mundo está equivocado.) ¿Cómo se supone que esto es superior al intuicionismo ético, incluso en la interpretación menos sofisticada del intuicionismo?

Es en este punto donde los antirrealistas expresarán su optimismo sobre lo mucho que la gente tiene realmente en común y sobre cuántas disputas morales resultarán depender de errores de hecho y de razonamiento.[31] Pero las razones para que las personas pensantes tengan sensibilidades fundamentalmente diferentes parecen al menos tan sólidas como las de que las personas pensantes tengan intuiciones fundamentalmente diferentes.

La mayoría de los antirrealistas querrían abrazar los métodos de la ética contemporánea, al igual que la mayoría de los éticos contemporáneos abrazan el antirrealismo. ¿Es esto coherente? La práctica actual de la filosofía moral incluye:

a) *La apelación omnipresente a las intuiciones* (con o sin usar la palabra "intuición"). Todo argumento moral emplea una o más intuiciones. Por ejemplo, el conocido argumento de Peter Singer a favor de una fuerte obligación de ayudar a los pobres: se basa en el "principio plausible" de que si podemos evitar algo malo sin sacrificar nada de importancia comparable, deberíamos hacerlo. No contento con esto, apela a un ejemplo hipotético en el que pasa por delante de un niño que se ahoga en un estanque poco profundo. Singer apela a nuestra intuición moral (¿cómo llamarla si no?) de que está obligado a ayudar al niño.[32] Como muestra este caso, se utilizan intuiciones tanto sobre principios generales como sobre casos particulares.

¿Cómo puede un antirrealista acomodar esto? El nihilista no puede tener nada que ver con esto, ya que piensa que los enunciados morales relevantes son todos falsos, a menos que, tal vez, proponga reformar

nuestro lenguaje moral, para que lleguemos a usarlo como los no-cognitivistas o los subjetivistas piensan que lo hacemos.

Para el no-cognitivista, el discurso moral sólo expresa nuestras actitudes. Para el subjetivista, informa de nuestras actitudes. De cualquier modo, parecería que la única tarea cognitiva para un filósofo moral es estudiar nuestros sentimientos, deseos y otras actitudes. Tal vez se dirá que los argumentos basados en intuiciones morales están simplemente al servicio de ese objetivo -Singer sólo está llamando nuestra atención sobre el hecho de que, de hecho, aprobamos los actos de prevenir-cosas-malas-sin-sacrificar-cosas-de-significado-comparable, incluyendo actos tales como sacar a los niños que se ahogan de estanques poco profundos. Tal vez su argumento no se base en intuiciones, sino en emociones morales.

Esto estaría bien, si no fuera por la segunda característica del discurso contemporáneo de la filosofía moral:

b) *Las críticas a los prejuicios*. A veces se argumenta que ciertas intuiciones o creencias morales no son fiables, ya que sólo se deben a nuestra cultura, nuestra religión o al simple egoísmo. Recordemos, por ejemplo, el argumento de Singer de que nuestras intuiciones sobre el infanticidio no son fiables, ya que sólo se deben a la programación cultural.[33] En otro lugar, ha sugerido que nuestra creencia en la superioridad de la especie humana se debe al egoísmo y a las enseñanzas religiosas que reflejan nuestro interés personal.[34]

Esto tiene sentido para un realista moral: si una creencia moral se debe a sesgos de un tipo que, en general, suele causar creencias falsas, entonces es probable que esa creencia no se corresponda con los hechos morales. Pero, ¿qué debería decir un antirrealista? La noción de "sesgo" presupone el realismo: un sesgo es un deseo, emoción u otro estado psicológico que tiende a distorsionar nuestras percepciones, a alejarnos

de la verdad. Eso presupone que existe una verdad independiente de nuestras percepciones. Para un relativista cultural, el argumento de Singer sobre el infanticidio es peor que un *non sequitur*: su premisa de que nuestras creencias sobre el infanticidio están causadas por el adoctrinamiento cultural demostraría que esas creencias son correctas, en lugar de socavarlas como supone Singer. Incluso para un subjetivista o un no cognitivista, nuestra cultura no es más que una fuente más de nuestras emociones morales; ¿con qué fundamento podrían considerarse "correctas" las reacciones morales causadas por ella.[35] Lo mismo se aplica al supuesto sesgo debido al interés propio: si el interés propio tiñe nuestras emociones morales, todo lo que demuestra es que el interés propio determina en parte lo que es moralmente correcto o bueno.

Tal vez el antirrealista diría que cuando nos demos cuenta de las causas de algunas de nuestras emociones morales, perderemos esas emociones. Quizá cuando nos demos cuenta de que nuestro horror ante el infanticidio está culturalmente programado, dejaremos de sentirlo. O tal vez no. Si las emociones morales estuvieran causadas por percepciones putativas de hechos morales, o si existiera alguna norma externa de "adecuación" de las emociones morales, quizá entonces perderíamos nuestras emociones morales al descubrir su fuente parroquial. Pero una vez que renunciamos a la idea de que existe una verdad moral independiente de nuestros sentimientos, ¿por qué el descubrimiento no sorprendente de que algunas de nuestras emociones morales están causadas por condicionamientos culturales debería socavar esos sentimientos?

Un tercer aspecto de nuestro discurso moral está estrechamente relacionado con el segundo:

c) *Las críticas a las intuiciones morales basadas en otras intuiciones*. Peter Singer, de nuevo, ha sugerido que "nuestros

> juicios intuitivos [sobre casos particulares] se basan en cosas que obviamente no tienen ninguna importancia moral". Por ejemplo, cree que nuestras intuiciones se inclinan hacia los que están físicamente más cerca de nosotros (de modo que pensamos que tenemos mayores obligaciones con las personas de nuestro barrio, por ejemplo, que con las de otros continentes). ¿Cómo sabe Singer que la proximidad física no es moralmente relevante? La respuesta parece ser que eso es "obvio" o "evidente"; en pocas palabras, es una intuición.[36]

El intuicionista puede entender esto: las intuiciones son una forma falible de acceso a la verdad moral. Podemos corregir las intuiciones erróneas apelando a otras más firmes y fiables. Pero, ¿qué debería decir un antirrealista?

Una propiedad "moralmente relevante" es aquella que afecta a si una cosa es buena, mala, correcta o incorrecta. Desde el punto de vista subjetivista o no cognitivista de la moralidad, esto sólo puede significar una propiedad que afecta a (un cierto subconjunto de) nuestras actitudes. Por lo tanto, si la proximidad física afecta realmente a nuestras actitudes morales, como dice Singer, eso sólo demuestra que la proximidad física es moralmente relevante.

Tal vez se podría argumentar que lo que es moral viene determinado, no por nuestras actitudes actuales, sino por las actitudes que tendríamos si llegáramos a estar plenamente informados y a ser racionales (eliminando todas las incoherencias y errores fácticos de nuestra mente, etc.; véase más arriba, sección 3.5). Desde un punto de vista realista, es plausible que una persona plenamente racional no tenga el sesgo hacia las personas físicamente cercanas que describe Singer, puesto que lo reconocería como un sesgo. Pero no parece haber nada incoherente en que tengamos esta actitud, ni depende de un error de hecho, ni Singer intenta argumentar que así sea. Entonces, ¿por qué debería un

antirrealista racional perder la actitud? El hecho de que el argumento de Singer sea considerado contundente por los éticos, en lugar de un obvio non sequitur, sugiere que los filósofos morales emplean ilícitamente una metodología intuicionista mientras pretenden rechazar el intuicionismo.

Ahora bien, aquí hay algunas cosas que normalmente no aparecen en la literatura ética:

d) *Las meras apelaciones a las prácticas comunes*. Los detractores de la noción de derechos de los animales no argumentan, por ejemplo, que como comer carne es lo habitual por aquí, es lo correcto. Los defensores de los derechos de los animales tampoco consideran que el maltrato generalizado de los animales refute sus propios puntos de vista morales.

e) *Meras apelaciones a las emociones y deseos del autor o de la audiencia*. Los que se oponen al aborto no dicen, por ejemplo, que el aborto es malo porque ellos o su público tienen sentimientos de simpatía hacia los pequeños fetos. Los defensores del aborto no argumentan que el aborto no es un asesinato porque el aborto no les hace sentir lo mismo que el asesinato.

f) *Retórica desnuda*. Los filósofos no intentan "convencer" a su audiencia de sus puntos de vista morales simplemente utilizando un lenguaje cargado de emociones, contando historias lacrimógenas o dirigiendo demandas a la audiencia.

Si alguna forma de subjetivismo, relativismo o no-cognitivismo es cierta, ¿por qué lo anterior no debería ser la práctica habitual?

Los subjetivistas y los no-cognitivistas nos aconsejarían ser más sutiles que simplemente apelar directamente a las actitudes actuales en todos los casos; dirían que tenemos que hacer algún razonamiento para asegurarnos de que nuestras actitudes son coherentes y no se basan en

errores de hecho. Tal vez en eso consiste el razonamiento moral. Sin embargo, esta sugerencia se parece muy poco al razonamiento moral real, y no es probable que consiga resolver muchas disputas. Recordemos la sugerencia que se ha comentado en la sección 3.3 de que la práctica de la esclavitud era criticable desde dentro de nuestra cultura porque no era coherente con nuestra costumbre de respetar a las personas en general: ¿por qué no debería el defensor de la esclavitud responder simplemente: "No, no tenemos una costumbre de respetar a todas las personas. Tenemos la costumbre de respetar a los blancos". Uno puede estar tentado a objetar que esto se basa en una distinción arbitraria, pero si lo que es moralmente arbitrario o significativo sólo está determinado por las costumbres, entonces el color de la piel es moralmente significativo, en muchas culturas. Compárese con la idea, tal vez apoyada por Singer, de que nuestra falta de ayuda a los pobres del Tercer Mundo es incompatible con nuestra simpatía general por otras personas: ¿por qué no habría que responder simplemente: "No, no tengo un sentimiento de simpatía por todas las demás personas. Tengo un sentimiento de simpatía por las personas cercanas, que pertenecen a mi grupo social"? Esas respuestas tendrían sentido si fuéramos subjetivistas o no cognitivistas -deberían tener al menos un lugar importante en el debate-, pero de hecho nadie da respuestas de ese tipo.

Reconozco que hay casos en los que actuamos de forma incoherente con nuestras costumbres: que unas pocas personas sorban de vez en cuando directamente del cuenco no refuta la afirmación de que nuestra costumbre es tomar la sopa con una cuchara; esas pocas personas simplemente están violando la costumbre. Del mismo modo, el hecho de que la gente maltrate ocasionalmente a los negros no refuta la afirmación de que tenemos una costumbre de respeto a las personas en general. Pero si los miembros de una sociedad maltratan intencionada y rutinariamente a los miembros de una determinada raza, no decimos que tengan dos

costumbres: una costumbre de respeto a todas las personas, más una costumbre de falta de respeto a una raza. Decimos que tienen una costumbre de respetar a los miembros de ciertas razas y no a otros. A las actitudes individuales se les aplica un punto similar al de las costumbres.

El antirrealista tiene, de hecho, métodos disponibles para resolver desacuerdos morales. El antirrealista debería emplear los métodos (d), (e) y/o (f). El intuicionista debería emplear (a), (b) y (c). Casi nadie encontraría más atractivo el enfoque antirrealista que el intuicionista.

Con esto concluye nuestro tratamiento de la objeción más popular al intuicionismo. La confianza generalizada en que el desacuerdo ético proporciona una refutación inmediata del intuicionismo parece descansar en poco más que caricaturas de la teoría, dobles estándares epistemológicos, y el desprecio de los fenómenos de desacuerdo y error fuera de la ética normativa. Los intuicionistas no sostienen que el error moral sea imposible, la ética no es ni mucho menos el único ámbito en el que encontramos desacuerdos persistentes, y los intuicionistas pueden considerar que los errores morales están causados por el mismo tipo de factores que provocan errores en otros ámbitos. Los intuicionistas tampoco pierden la oportunidad de intentar reducir los desacuerdos morales; pueden respaldar las mismas estrategias argumentativas empleadas actualmente por los filósofos morales, incluidas las estrategias diseñadas para identificar qué intuiciones tienen más probabilidades de estar equivocadas. Los intuicionistas no pueden proInir una resolución rápida de la mayoría de las controversias morales, pero tampoco puede hacerlo nadie más. Por último, los argumentos contra el intuicionismo basados en el desacuerdo podrían esgrimirse con al menos la misma fuerza contra otras teorías metaéticas, pero sólo contra el intuicionismo se han considerado persuasivos. Esto sugiere que el argumento del desacuerdo ha parecido plausible sólo porque los

filósofos ya habían rechazado el intuicionismo por otras razones.

Notas

1. Brandt (1996, p. 5); lo que propone ignorar es la idea de que las intuiciones son guías fiables para las cuestiones morales sobre casos concretos. El no naturalismo se descarta en unas pocas frases en la página anterior. Entre los que ofrecen teorías generales de la naturaleza de la moral e ignoran el intuicionismo están Joyce (2001), Blackburn (1993) y Harman (1977). El excelente y ampliamente utilizado libro de texto de James Rachels (2003) no contiene ninguna mención al intuicionismo, aunque incluye secciones sobre otras teorías metaéticas. El libro de texto de Liszka opina que "los defectos de este enfoque son obvios" antes de relatar con ligereza el problema del desacuerdo (2002, p. 312). Jackson (2003, p. 562) observa: "Parece que sabemos lo suficiente sobre cómo es nuestro mundo como para saber que no existe tal propiedad" como la propiedad no natural de la bondad de Moore; no da más detalles. Jackson, Pettit y Smith (2004, p. 221) tachan el no naturalismo de Moore de "ridículo", también sin más detalles.

Sin embargo, el intuicionismo ha ganado un renovado interés en algunos sectores en los últimos años (Stratton-Lake 2002b; Audi 1997 y 2004; Dancy 1993; Shafer-Landau 2003).

2. Véase Howard y Dunaif-Hattis (1992, pp. 314-15, 461, 573). Algunos antropólogos niegan la autenticidad de la mayoría de los supuestos casos de canibalismo; véase Arens (1979) y Howard y Dunaif-Hattis (1992, p. 489).

3. Rachels (2003, pp. 24-5).

4. Hare (1999, pp. 5-6) argumenta que, dado que las intuiciones pueden diferir, el intuicionismo conlleva el relativismo. Pero véase la nota 5 más abajo.

5. Esta es la noción de objetividad de Hare (1999, p. 11).

6. Mackie (1977, pp. 36--8); Hudson (1967, pp. 58-60); Layman (1991, pp. 179-80).

7. Nowe II-Smith (1954, pp. 46-7). Hudson (1967, p. 58), Liszka (2002, p. 312), Popper (2001, pp. 49-50) y Tara Smith (2000, pp. 25-6) atribuyen la pretensión de infalibilidad al intuicionismo. Smith atribuye explícitamente la afirmación de la irrelevancia del razonamiento, mientras que los demás parecen implicarla. La afirmación (iii) no se menciona explícitamente, pero parece requerida por el argumento. Brandt

(1944, p. 477) implica que las intuiciones intelectuales deben ser infalibles, bajo condiciones apropiadas.

8. Véase el capítulo 5, nota 12. No puedo encontrar un apoyo claro a la afirmación de infalibilidad general en toda la literatura intuicionista.

9. Clarke ([1705]1964, p. 6). Citado en Nowell-Smith (1954, p. 47). He abreviado la cita.

10. Clarke ([1705] 1964, pp. 4-5). Sus otros ejemplos son similares, siendo generalmente del tipo: es mejor ayudar a la gente que causar miseria y destrucción sin razón.

11. Brandt (1944, p. 483).

12. Arens (1979); Mead (2001); Freeman (1983; 1999). 13. Véase Gardner (1998) para una revisión de la controversia sobre el canibalismo. Véase Hellman (1998, capítulo 10) para una revisión de la controversia Mead-Freeman. Hellman también revisa otras nueve grandes controversias en la ciencia.

l3. Murray (1984).

14. Véase Newport y Strausberg (2001) sobre la percepción extrasensorial y la astrología y Gallup y Newport (1991, pp. 138, 141) sobre la percepción extrasensorial, la astrología y las experiencias telepáticas. Newport y Strausberg informan de que, en 2001, el 36% de los estadounidenses creía en la telepatía, aunque Newport y Strausberg no detallan cuántos creen haber tenido realmente una experiencia telepática.

15. Nowell-Smith (1954, p. 46); Hudson (1967, pp. 58-60).

16. Véase mi (n.d.) sobre la irracionalidad y el sesgo.

17. Esta es la posición de la "Tierra joven" que sostienen muchos creacionistas, como Thompson (1999). La cifra de 6000 se obtiene sumando las edades de las personas mencionadas en la Biblia, especialmente en el Génesis.

18. Compárese con Shafer-Landau (2003, p. 220).

19. Nowell-Smith (1954, pp. 44-6).

20. Sterling (1994, p. l30) y Brink (1984, p. 117) hacen observaciones similares. Para ejemplos de argumentos racionales sobre el aborto, véase Thomson (1971), Tooley (1972) y Marquis (1989). Aunque estos argumentos no han logrado resolver la disputa a

satisfacción de todos, son un paso en la dirección correcta.

21. Brink (1984, p. 116) argumenta de forma similar, aunque parece pensar erróneamente que el punto requiere una teoría cohercntista de la justificación. Mi punto de vista se asemeja a lo que Haack (1983) denomina "foundherentismo", pero nótese que no asumo, como Haack, que todas las creencias fundacionales sean falibles. Compara las nociones de Bonjour (1985, pp. 26-9) de "fundacionalismo moderado" y "débil", que permiten que la coherencia aumente o disminuya la justificación de las creencias fundacionales.

22. El punto no depende de la deshonestidad de los testigos. Los estudios han revelado que la fiabilidad del testimonio de los testigos oculares está muy sobrevalorada, los testigos pueden estar muy seguros incluso cuando se equivocan, y muchas personas condenadas sobre la base de tales testimonios han sido exoneradas posteriormente. Véase una revisión en Wells y Olson (2003).

23. Sinnott-Armstrong (1996, pp. 26-7); Layman (1991, pp. 179-80); Ayer (1952, p. 106).

24. Nowell-Smith (1954, p. 47) pretende hablar en nombre de los intuicionistas: "Los que no están de acuerdo con nosotros en una cuestión moral deben ser insinceros. Lo que necesitan es la fuerza, no el argumento".

25. Singer (1993, pp. 169-73). Está de acuerdo en que "ciertamente deberíamos poner condiciones muy estrictas al infanticidio permisible" (p. 173).

26. Este último principio es criticado por Stuart Rachels (1998), pero no me convence.

27. Stuart Rachels (2003) enumera nueve formas de argumentar las afirmaciones sobre el valor intrínseco, siete de las cuales parecen viables. Ninguna de las viables entra en conflicto con mi tesis de que la intuición es la fuente última de justificación de todas las afirmaciones éticas.

28. Las observaciones en este sentido son comunes; véase Rawls (1971, p. 21); Timmons (1999, p. 7); Smith (1994, pp. 201-2).

29. Esta es la explicación de Stevenson (1963, pp. 26-7) sobre el desacuerdo moral.

30. Joyce (2001, capítulo 6).

31. Véase Singer (1993, p. 327); Hare (1999, pp. 24, 58-9); Ruse (1998, p. 255).

32. Singer (1993, pp. 229-30). Hay más discusión, por supuesto, pero la apelación a la intuición es esencial.

33. Ver arriba, pp. 145-7.

34. Singer y Posner (2001, carta inicial). Para más sugerencias sobre nuestros prejuicios, véase Singer (1974, p. 516).

35. Entre los antirrealistas, sólo Scanlon parece reconocer este punto: "Si, no obstante, se nos acusa de que nuestros juicios sobre las razones para la acción en su conjunto se deben "meramente" a la habituación, ¿qué concepción de las razones respalda este "meramente"? ¿"Simplemente" en contraposición a qué? Si el contenido de los juicios sobre las razones sugiriera que se supone que nos revelan los hechos sobre alguna realidad "externa", entonces esta acusación podría al menos ganar un punto de vista" (1998, p. 72).

36. La cita es de Singer (1999, p. 316), refiriéndose con aprobación a los argumentos de Peter Unger. Unger (1996, p. 101) da varios ejemplos de afirmaciones morales obviamente falsas de las que cree que dependen nuestras intuiciones, aunque no afirma que las in tuiciones morales sobre los casos no sean fiables en general (p. 10). Singer (1999, p. 317) declara: "Estos factores simplemente no pueden ser moralmente significativos" - fin del argumento.

Véase también Singer (1993, pp. 232, 254) sobre el sesgo de proximidad física. Singer (1974, pp. 516-17; 1999, p. 270) recomienda que nos basemos en "axiomas evidentes" en lugar de en intuiciones sobre los casos.

7

Razones prácticas

Muchos antirrealistas han afirmado que si el realismo moral fuera cierto, no habría ninguna razón para actuar moralmente. En este capítulo, sostengo en cambio que sólo el intuicionista tiene una explicación plausible de nuestras razones para ser morales.

7.1 El argumento humeano contra el realismo

¿Qué razones tenemos para comportarnos moralmente? Hay al menos dos interpretaciones de esta pregunta, porque hay dos sentidos de "razón": hay razones motivadoras y razones normativas. Las razones motivadoras (o "motivos") son el tipo de cosas que explican por qué la gente hace lo que hace. Tener una razón motivadora para hacer *A* no implica que uno vaya a hacer *A*; sólo significa que uno tiene al menos cierta tendencia o tentación de hacer *A*, dada la oportunidad. Por ejemplo, mi hambre, combinada con mi creencia de que el sándwich que tengo delante es comida, constituye un motivo para que me coma el sándwich.

Los motivos normativos son el tipo de cosas que explican por qué la gente debe hacer ciertas cosas, desde el punto de vista de la razón.[1] Una

razón normativa para realizar una acción es una consideración a favor de la misma, aunque no necesariamente una consideración decisiva. Al igual que las razones motivadoras, las razones normativas pueden entrar en conflicto. Lo que es racional hacer es lo que uno tiene más razones normativas para hacer. Por ejemplo, el hecho de que necesite comida, junto con el hecho de que el bocadillo que tengo delante sea comida, constituye una razón normativa para que me coma el bocadillo.

Normalmente tenemos razones normativas y razones motivadoras para las mismas cosas. En el ejemplo anterior, tengo tanto una razón motivadora como una razón normativa para comerme el bocadillo. Pero a menudo la fuerza de nuestros motivos no se corresponde con la fuerza de nuestras razones normativas. Supongamos que puedo elegir entre un nutritivo sándwich vegetal y un delicioso helado. Puede que tenga mejores razones para comerme el sándwich vegetariano, pero puede que tenga más motivación para comerme el helado. El placer gustativo anticipado del helado es una justificación para elegirlo (en la medida en que el placer es un bien), pero esta justificación puede verse superada racionalmente por consideraciones de salud a largo plazo. Estas consideraciones de salud a largo plazo producen en mí cierta tendencia a elegir las verduras, pero esta tendencia puede verse superada, desde el punto de vista motivacional, por el deseo de placer azucarado y, por tanto, puedo hacer una elección irracional.

Los tres principios siguientes sobre las razones para actuar están ampliamente aceptados en filosofía; los dos primeros son también supuestos estándar en economía:

i) "La concepción humeana de las razones" (en adelante, "la concepción humeana"): Todas las razones para actuar derivan de deseos preexistentes. En concreto:

a) Tienes una razón que te motiva a hacer *A* si y sólo si hay algo que

deseas por sí mismo y que crees que *A* lograría o podría lograr. La fuerza de tu razón para hacer *A* es proporcional a la fuerza de ese deseo.

b) Tienes una razón normativa para hacer *A* si, y sólo si, hay algo que deseas por su propio bien y que tienes razones para creer que *A* lograría o podría lograr.[2]

ii) La brecha creencia/deseo: las creencias no implican deseos: es posible creer que se obtiene cualquier hecho objetivo dado, mientras se carece de cualquier deseo dado.

iii) El magnetismo de los valores:

a) Si uno acepta un valor moral, tiene necesariamente una razón que motiva la acción. Por ejemplo, si uno considera que robar es malo, entonces debe tener al menos alguna motivación para evitar el robo.

b) Los hechos morales, si los hay, proporcionan necesariamente razones normativas para la acción. Si robar es malo, o al menos si uno tiene razones para creer que es malo, entonces uno tiene una razón normativa para evitar robar.

Se suele pensar que estos principios refutan el realismo moral. ¿Cómo es eso? Hay dos versiones del argumento:[3]

Primera versión:

1. Los motivos de la acción requieren deseos. (La concepción humeana).

2. Ninguna creencia sobre un hecho objetivo implica la presencia de ningún deseo. (La brecha creencia/deseo).

3. Por lo tanto, ninguna creencia sobre un hecho objetivo, por sí misma, implica la presencia de ningún motivo para la acción. (De 1, 2.)

4. Las actitudes morales (las actitudes expresadas por las afirmaciones morales) implican la presencia de motivos para la acción. (El magnetismo de los valores.)

5. Por lo tanto, las actitudes morales no son creencias sobre hechos objetivos. (De 3,4.)

Segunda versión:

6. Las razones normativas para la acción requieren deseos. (La concepción humeana).

7. Ningún hecho objetivo implica la presencia de un deseo. (Porque los hechos objetivos son hechos del mundo externo, al margen de nuestros deseos y actitudes).

8. Ningún hecho objetivo, por sí mismo, implica la presencia de una razón normativa para la acción. (De 6, 7.)

9. Los hechos morales, si los hay, implican la presencia de razones normativas para la acción. (El magnetismo de los valores.)

10. Por lo tanto, no hay hechos morales objetivos. (De 8, 9.)

En resumen, las afirmaciones morales enuncian razones para la acción. Las razones para la acción, de cualquier tipo, requieren la existencia de deseos. Pero ningún hecho objetivo, o creencia sobre un hecho objetivo, puede implicar por sí mismo la existencia de ningún deseo. Por lo tanto, las afirmaciones morales no se limitan a informar de hechos objetivos; deben contener algún elemento subjetivo o no cognitivo.

Un realista moral puede responder de tres maneras: rechazar la concepción humeana, rechazar la brecha creencia/deseo o rechazar el magnetismo de los valores. Cada camino ha sido tomado por al menos algunos filósofos.[4] Criticaré la Concepción Humeana y propondré una concepción alternativa de las razones.

7.2 La conexión entre razones motivadoras y normativas

Una respuesta natural a la primera versión del argumento humeano es cuestionar (4). Si considero que robar está mal, entonces quizás se deduce que debería estar, o que creo que debería estar, motivado para evitar robar. Pero no se deduce obviamente que de hecho esté motivado para evitar robar. Podría ser una persona inmoral, incluso según mis propios criterios, es decir, podría no estar motivado por lo que considero moralmente correcto. Por razones similares, podríamos dudar de (6) en su segunda versión. Si tengo una razón normativa para ayudar a Jimmy con sus deberes, quizás se deduce que debería desear ayudar a Jimmy, o que desearía hacerlo si fuera plenamente racional, pero no se deduce obviamente que de hecho tenga algún deseo que se satisfaría ayudando a Jimmy. Hay razones para sospechar que el argumento se basa en un equívoco: la tesis humeana -que las razones para la acción requieren deseos- sólo es plausible si tomamos "razones" como "razones motivadoras". Pero la premisa del magnetismo de los valores -que las afirmaciones morales establecen razones para la acción- sólo es plausible si entendemos por "razones" "razones normativas". Es decir, es plausible que nuestra motivación para hacer algo requiera que tengamos un deseo adecuadamente relacionado con la acción, pero no que nuestra razón normativa para hacer algo requiera tal deseo; y es plausible que nuestra justificación de que algo es bueno nos dé una razón normativa para perseguirlo, pero no que realmente nos motive a perseguirlo. Pero el argumento requiere que tanto la concepción humeana como el magnetismo de los valores sean verdaderos para el mismo sentido de "razón".

Sin embargo, esa no es mi respuesta definitiva al argumento. Planteo este punto para mostrar por qué un defensor del argumento humeano

debería querer defender alguna conexión fuerte entre las razones motivadoras y las normativas, aproximadamente en el sentido de que los dos tipos de razones siempre van juntos. Afortunadamente para ellos, pueden hacerlo mediante al menos dos argumentos plausibles.

En primer lugar, un argumento basado en el principio del "deber implica poder": decir que uno debe hacer *A* implica que puede hacer *A* por voluntad propia.[5] Esto implica que uno puede elegir libremente hacer *A*, lo que a su vez implica que uno puede estar motivado para hacer *A*; una persona no puede elegir libremente hacer algo sin tener ninguna motivación para hacerlo. Ahora bien, decir que tienes una razón normativa para hacer *A* implica que deberías hacer *A* (siempre que no haya consideraciones contrarias). De ello se deduce que eres capaz de estar motivado para hacer *A* (siempre que no haya consideraciones compensatorias). Pero si no tuvieras ningún motivo que te pareciera útil para hacer *A*, entonces, puesto que tampoco puedes elegir qué motivos tienes, serías incapaz de estar motivado para hacer *A*. Por lo tanto, la afirmación de que tienes una razón normativa para hacer *A* implica que tienes alguna motivación disponible para hacer *A*.[6] En resumen:

1. Si *S* tiene una razón normativa para hacer *A*, entonces *S* debe hacer *A*, siempre que no haya razones en contra.

2. Si *S* debe hacer *A*, entonces *S* puede hacer *A* libremente.

3. Si *S* puede hacer libremente *A*, entonces *S* puede estar motivado para hacer *A*.

4. Por lo tanto, si *S* tiene una razón normativa para hacer *A*, entonces *S* puede estar motivado para hacer *A*, siempre que no haya razones compensatorias. (De 1,2,3.)

5. Si *S* no tiene una razón motivadora disponible para hacer *A*, entonces *S* no puede estar motivado para hacer *A*, incluso si no hay razones

compensatorias contra *A*.

6. Por lo tanto, si *S* tiene una razón normativa para hacer *A*, entonces *S* tiene una razón motivadora disponible para hacer *A*. (De 4, 5.)

Para ilustrar, supongamos que alguien me dice que tengo una razón normativa para donar 500 dólares a Oxfam. Supongamos, sin embargo, que estoy tan falto de preocupación por mis semejantes que no tengo ningún deseo de ayudar a los pobres, ni ningún otro deseo que se satisfaga donando ese dinero. Supongamos que después de una profunda y correcta deliberación, me doy cuenta de estos hechos. Además, no puedo evitar nada de esto; no puedo evitar el hecho de que no me importan los pobres, ni el hecho de que sé que donar 500 dólares a Oxfam no satisface ninguno de mis deseos. En esta situación, no puedo estar motivado para donar el dinero. Por lo tanto, simplemente no puedo donar el dinero por mi propia voluntad. Y, por tanto, es falso decir que debo donar el dinero. A la inversa, para que sea cierto que debo donar el dinero, uno de los supuestos anteriores tendría que ser falso; tendría que tener algún deseo que pensara que la donación del dinero podría satisfacer.

Este es un argumento que podría esgrimir un humeano, para establecer la conexión entre las razones motivadoras y las normativas, y así reforzar la afirmación de que "*S* debería hacer *A*" implica que *S* tiene algún tipo de deseo.

Un segundo argumento a favor de esta conexión lo llamo el Argumento de la Interpretación. Nuestra forma habitual de interpretar el comportamiento de otras personas (incluidas sus declaraciones) se basa en un supuesto de racionalidad: cuando se intenta averiguar lo que otra persona cree, desea, etc., se observa su comportamiento y se le atribuyen las creencias, deseos, etc. que harían que ese comportamiento tuviera más sentido. Por ejemplo, si alguien come mucho chocolate, deduzco

que le gusta el chocolate y que, por tanto, está emprendiendo el camino correcto para satisfacer sus preferencias. Si por el contrario asumiera que es irracional, podría hipotetizar que odia el chocolate y que por eso lo come. Algo parecido ocurre con la atribución de creencias a los demás: Asumo que mi amiga que come chocolate cree que está comiendo chocolate, porque eso es lo que tendría sentido que creyera, dadas sus circunstancias reales. Si la considerara irracional, podría hipotetizar que cree que está nadando en el Canal de la Mancha, y que le gusta nadar en el Canal, así que por eso se come el chocolate. Las hipótesis que presuponen la irracionalidad pueden "explicar" el comportamiento seguro, pero las consideramos, con razón, ridículas a primera vista. La cuestión es que tenemos que reconocer que los demás son, al menos en general, racionales para poder elegir una explicación adecuada de su comportamiento entre las muchas posibles.

Algo parecido a este principio está ampliamente aceptado bajo el nombre de "Principio de Caridad".[7] El principio no consiste en que todas las acciones o creencias sean racionales, sino sólo en que debemos interpretar que los demás son, en general, racionales, si es posible. A veces, las declaraciones y el comportamiento de las personas demuestran que son irracionales en ciertos aspectos. El Principio de Caridad no niega esto, pero sostiene que la racionalidad es una suposición por defecto, y que no debemos atribuir más irracionalidad de la necesaria. No se trata simplemente de cómo interpretamos a los demás, sino de cuál es la forma correcta de interpretar a los demás. Obsérvese también que el principio no es empírico: no se trata de que, observando a las personas durante mucho tiempo, hayamos descubierto que son en general racionales. Más bien, la suposición de racionalidad general es algo que aportamos a nuestras observaciones, como una restricción a su interpretación. No observamos directamente las creencias y los deseos de los demás, por lo que no tendríamos forma de averiguar si las

personas son en general racionales, sin tener primero alguna forma de inferir sus estados mentales a partir de su comportamiento. Ese tipo de inferencia es precisamente lo que es la "interpretación". Dado que la suposición de racionalidad es necesaria para interpretar el comportamiento, la suposición de racionalidad no puede ser probada observando el comportamiento de las personas. A lo sumo, lo que se puede descubrir empíricamente es que las personas exhiben un comportamiento que puede ser interpretado racionalmente, pero no podemos determinar elnpíricamente que las interpretaciones de las personas como racionales son *pro* tanto mejores que las interpretaciones competidoras de las personas como irracionales.

El principio de la caridad es casi universalmente aceptado. Pero, ¿cuál es su fundamento? Parece que el principio sólo podría ser correcto si estuviéramos justificados en la creencia, en general, de que las personas son en general racionales. Pero, por las razones que acabamos de exponer, esta creencia no puede ser justificada por la observación. Por lo tanto, tiene que ser a priori. ¿Cómo puede ser eso? Aquí hay una respuesta natural: porque es una verdad necesaria que siempre que se tiene una razón normativa para la acción, se tiene una razón motivadora para la acción. Supongamos lo contrario: supongamos, es decir, que uno puede tener una razón normativa para hacer *A* pero ningún motivo para hacer *A*. Si esto es posible, entonces se ve que también sería posible que una persona careciera de motivos, en general, para hacer las cosas que tenía razones normativas para hacer. Tal persona sería generalmente irracional. Pero si es posible que una persona sea generalmente irracional, entonces ¿cómo podríamos saber a priori que las personas son en general racionales?

El argumento anterior puede ser más débil en su primera etapa, en la que interpreté que el Principio de Caridad implica que debemos

interpretar que las personas son en general racionales. Tal vez, en cambio, sólo debamos interpretar que las personas son racionales en ciertos aspectos, lo que deja abierta la posibilidad de que descubramos otros aspectos amplios y sistemáticos en los que las personas no son racionales. Sin embargo, no voy a seguir esta línea. Me parece plausible que las razones normativas tengan el tipo de "magnetismo" que plantea el argumento humeano. La idea no es que todos los seres sean movidos por todas las razones normativas. La idea es que si un ser puede y reconoce una razón normativa, entonces el ser debe tener al menos alguna tendencia a ser movido por esa razón. Pero puede haber muchos seres que no reconozcan muchas razones normativas.

Hasta ahora he tratado de exponer el caso humeano con la mayor firmeza posible: parece que las afirmaciones valorativas implican la existencia de razones normativas para la acción; esto, a su vez, implica o sugiere fuertemente la existencia de razones motivadoras para la acción; y eso implica la existencia de deseos que están apropiadamente relacionados con las acciones en cuestión. Pero ningún hecho objetivo puede implicar la existencia de un deseo; por lo tanto, las afirmaciones evaluativas no son afirmaciones objetivas y fácticas. Donde cuestiono este argumento es en su suposición de que toda motivación proviene del deseo.

7.3 Una concepción racionalista de la motivación

Existen al menos cuatro tipos de motivaciones a las que estamos sometidos:

i) *Apetito*: Los apetitos incluyen cosas como el hambre, la sed y la lujuria. Se trata de deseos simples, a menudo instintivos, cuya satisfacción suele ir acompañada de placer físico (es decir,

sensorial[8]), y/o cuya frustración va acompañada de dolor o malestar físico.

ii) *Emociones*: Los deseos emocionales, como los provocados por la ira, el miedo y el amor, se diferencian de los apetitos en que su satisfacción o frustración se asocia normalmente a un tipo diferente (emocional) de disfrute o sufrimiento, en lugar de a un placer o dolor sensorial. Además, los deseos emocionales requieren un tipo de cognición más sofisticado que los apetitos. Enfadarse con alguien implica percibir que algo que ha hecho es malo en cierto modo; temer algo implica percibirlo como peligroso; amar a alguien implica percibir su existencia y la relación con él como algo valioso; y así sucesivamente. Para los apetitos no se requieren percepciones tan sofisticadas. Por su parte, las emociones pueden tener objetos mucho más sofisticados; uno puede enfadarse por un hecho que requiere conceptos bastante abstractos para ser consciente, por ejemplo, el hecho de que un estudiante haya plagiado un trabajo. Nuestras emociones también suelen tener efectos a largo plazo: ese enfado puede provocar que se le ponga una nota baja al estudiante, meses después. Los apetitos tienen objetos más sencillos y efectos a más corto plazo.

iii) *Prudencia*: La prudencia es la motivación para perseguir (o evitar) una cosa porque se reconoce que favorece (o retrasa) los propios intereses. La prudencia difiere de los deseos apetitivos y emocionales, y a menudo se opone a ellos. Para volver a un ejemplo anterior, mi apetito por el helado puede entrar en conflicto con mi motivo prudencial de mantener una buena salud. Del mismo modo, un enfado excesivo con alguien puede tentarme a pegarle, mientras que la prudencia me pide que me contenga.

iv) *Razones imparciales*: A veces nos motiva a actuar de determinada manera nuestro reconocimiento de lo que es bueno, justo, honesto, etc. Estas razones son imparciales, en el sentido de

que no apelan a nuestros propios intereses o a nuestras preferencias personales. Este tipo de motivo puede entrar en conflicto con cualquiera de los tres primeros; por ejemplo, un hombre puede desear engañar a su mujer, pero creer que eso sería infiel (y por tanto malo). O un juez puede considerar la posibilidad de aceptar un soborno, lo que promovería sus propios intereses, pero creer que esto sería deshonesto (y por lo tanto malo).

Estos parecen ser los principales tipos de motivos. Hay varias cuestiones interesantes sobre ellos. En primer lugar, ¿cuál es el objeto de un deseo apetitivo? Es decir, cuando uno tiene apetito, ¿qué es lo que quiere? Una opinión tradicional es que se trata del placer físico o de evitar el dolor físico, lo que ha llevado a algunos pensadores a generalizar que todo deseo se reduce al deseo de placer y a la ausencia de dolor.[9] Sin embargo, esto es un error. Uno puede, por supuesto, desear placer, pero no es en esto en lo que consisten los apetitos típicos, aunque el placer normalmente acompaña a la satisfacción de esos apetitos. Cuando uno tiene hambre, lo que quiere es comida. También puede desear placer gustativo, pero el hambre no consiste en ese deseo, sino en el deseo de comer. Para ilustrarlo, supongamos que tienes hambre. Unos científicos locos estimulan tu cerebro para que tengas una alucinación increíblemente realista de comerte un filete de salmón, y para que al mismo tiempo dejes de sentir hambre. En este caso, yo diría que tu deseo (el hambre) desapareció sin quedar satisfecho: no conseguiste lo que querías, aunque te hayan engañado haciéndote creer que sí. Nuestros deseos suelen estar centrados en el mundo, no en nuestra mente.[10]

Del mismo modo, ¿cuáles son los objetos de los deseos emocionales? Si uno tiene la tentación de considerar los apetitos típicos como deseos de placer físico, podría tener la misma tentación de tratar los deseos emocionales como deseos de satisfacción emocional. Así, uno podría considerar que el amor de un hombre por sus hijos da lugar

principalmente a un deseo de satisfacción emocional por verlos bien. Pero, en realidad, lo que suscita el amor es el deseo de que los seres queridos estén mejor. Por eso, por ejemplo, tomar medidas para que los hijos estén bien después de su muerte es una manifestación de amor.

Para ilustrarlo, imagina que te ofrecen la posibilidad de vender a tus hijos como esclavos por 10.000 dólares. Los hijos serán desgraciados durante el resto de sus vidas. Sin embargo, como parte del trato, usted recibirá una sesión de hipnosis que le hará olvidar que los niños fueron vendidos como esclavos y, además, inducirá en usted una ilusión permanente de que los niños son extremadamente felices y exitosos. De este modo, tendrá la satisfacción de pensar que sus hijos están bien. Suponiendo que usted quiera mucho a sus hijos, ¿aceptaría el trato? Si su respuesta es "no", entonces su principal deseo no es la satisfacción emocional.[11]

¿Dónde encaja el deseo de placer en mi taxonomía de motivaciones? Creo que encaja en la prudencia, en el sentido ciertamente amplio en que he definido la "prudencia". El deseo de asegurar los placeres para uno mismo es un deseo de promover los propios intereses de una manera determinada. También es posible que simplemente tengamos un apetito por el placer.

Esto nos lleva a otra teoría psicológica simplista pero común: la teoría de que todas las acciones están motivadas por el interés propio. Esta teoría se enfrenta a la misma objeción que la anterior. Nos mueve una variedad de deseos diferentes -deseos de comida, sexo, poder, riqueza, amor, fama, etc.- la mayoría de los cuales son por cosas lógicamente independientes del avance de nuestros intereses. No es necesario hacer una búsqueda, ni siquiera ejemplos particularmente nobles, para refutar el egoísmo psicológico. Consideremos un deseo, provocado por la ira, de golpear a alguien: no es necesario que uno

considere que esa acción es de alguna manera en su interés para sentir el deseo, y uno puede muy bien considerarlo muy contrario a sus intereses.[12]

Estas observaciones socavan dos teorías humeanas particulares de las razones: la opinión de que todas las acciones están motivadas por el deseo de asegurar el placer y evitar el dolor, y la opinión de que todas las acciones están motivadas por el deseo de promover los propios intereses. Pero dejan intacto el marco humeano más general, según el cual todas las acciones están motivadas por deseos de un mismo tipo.

¿Qué lugar ocupa la moral en esta taxonomía? Considero que los principios morales son una especie de razones imparciales. Así, las consideraciones sobre lo que es bueno, justo, honesto, etc., son tanto razones morales como razones imparciales para la acción. Puede haber otras razones imparciales que no sean morales (tal vez el hecho de que un determinado procedimiento sea la forma más eficiente de producir lápices sea una razón imparcial, pero no una razón moral, para que uno haga que la fábrica de lápices fabrique lápices de esa manera), pero no las discutiré aquí.

Siempre tenemos una razón imparcial para hacer lo que sabemos que es moralmente correcto. Pero esto no quiere decir que esa razón sea la única, ni siquiera la principal, que tenemos para actuar moralmente. También tenemos muchas razones no morales para ser morales. Pensemos, por ejemplo, en una persona que hace lo que cree que es correcto, no sólo porque es correcto, sino porque cree que Dios castigará a los malhechores; o en una persona que se abstiene de robar, no sólo porque reconoce que el robo está mal, sino porque robar le hace sentir mal. Estos serían, respectivamente, motivos prudenciales y emocionales para hacer lo que es correcto.

¿Cómo encaja la noción de "deseo" en la taxonomía de los motivos? En el lenguaje ordinario, los apetitos y los motivos emocionales se denominan habitualmente "deseos", mientras que los motivos prudenciales e imparciales no lo son. Imaginemos a una estudiante que sabe que tiene un examen mañana por la mañana, para el que no ha estudiado. Una amiga la invita a una fiesta. El diálogo podría ser el siguiente:

Amigo: ¿Quieres ir a una fiesta esta noche?

Estudiante: Me apetece mucho ir, pero esta noche tengo que estudiar.

Amigo: ¿Dices que prefieres estudiar a salir de fiesta?

Estudiante: Claro que no quiero estudiar; prefiero estar de fiesta. Pero realmente necesito estudiar.

La interpretación natural es que el estudiante ha descrito correctamente sus motivaciones: tiene un deseo emocional de salir de fiesta, no tiene ganas de estudiar, pero reconoce -es decir, cree- que le conviene mucho estudiar (lo "necesita"). Esta última motivación, la prudencial, puede prevalecer sobre su deseo.

O imaginemos a un juez al que se le ha ofrecido un soborno para que absuelva a un criminal. El diálogo podría ser el siguiente:

Juez: No aceptaré su soborno.

Criminal: ¿Cuál es el problema, no quiere 100.000 dólares?

Juez: Por supuesto que quiero 100.000 dólares. Pero absolverte sería un error.

¿Así que está diciendo que quiere enviarme a la cárcel más que tener 100.000 dólares? Debe de odiarme de verdad.

Juez: No importa lo que sienta por ti. Usted cometió un delito, y los delincuentes deben ser castigados, y mi trabajo es velar por ello.

Criminal: Apuesto a que mi dinero podría comprarle mucho más placer que la satisfacción que obtendría al verme castigado.

Juez: Usted no lo entiende.

La interpretación natural es que el juez reconoce que aceptar el soborno promovería mejor sus propios intereses, satisfaría sus deseos y aumentaría su disfrute, pero cree que estaría mal. Su creencia moral prevalece sobre sus intereses personales.

¿Qué diría el humeano? Que en el primer ejemplo, la estudiante quiere sacar una buena nota en el examen, más que irse de fiesta, y que en el segundo ejemplo, el juez quiere que el criminal sea castigado más que los 100.000 dólares. ¿Es esto cierto? No en ningún sentido no trivial. En inglés, "desire" parece referirse a menudo a apetitos y emociones, como ilustran los ejemplos anteriores. En ese sentido, la afirmación del humeano es falsa. El estudiante puede sentir más deseo de irse de fiesta que de sacar una buena nota en el examen, pero aun así reconocer que le convendría más estudiar. El juez puede sentir más deseos de llevarse los 100.000 dólares que de castigar al delincuente, pero reconocer que la justicia le obliga a castigar al delincuente. Introspectivamente, simplemente no hay pruebas para la afirmación humeana.

En otro sentido, sin embargo, un "deseo" es simplemente cualquier motivación, sea del tipo que sea.[13] En este sentido, es trivial que en mis ejemplos el estudiante desee estudiar y el juez desee rechazar el soborno. En esta interpretación, la tesis fundamental humeana de que los motivos para la acción requieren deseos no es más que la siguiente tautología, con la que estará de acuerdo incluso el racionalista más acérrimo: Los

motivos para la acción requieren motivos para la acción. Pero seguramente no se puede inferir de ello ninguna conclusión interesante. En particular, en la concepción actual del deseo, ya no se puede suponer que las creencias no sean deseos; si una creencia motiva la acción, entonces sería ipso (acto un "deseo" en el sentido actual. Y simplemente tendríamos que distinguir dos tipos de "deseos": los que están constituidos por creencias y los que están constituidos por apetitos o emociones.

Supongo que el humeano no quiere trivializar su tesis de esta manera; la mayoría de los humeanos dirán que tenemos alguna concepción independiente de "deseo", como un tipo de estado mental identificable introspectivamente, sobre el cual no es una cuestión abierta si un deseo puede ser una creencia.[14] Pero en la introspección, los únicos estados mentales de este tipo que puedo encontrar que podrían llamarse "deseos" son los apetitos y las emociones. Y no es discutible si siempre actúo según mis apetitos y emociones más fuertes en ese momento -sé que, afortunadamente, no lo hago-.

7.4 ¿Por qué creer en la concepción humeana?

7.4.1 El atractivo intuitivo de la concepción humeana

¿Por qué tantos filósofos han adoptado la concepción humeana de las razones? La mayoría parece simplemente considerarla intuitivamente obvia: es obvio que las razones para la acción requieren objetivos, y es obvio que sólo nuestros deseos, y no meras creencias sobre el mundo exterior a nosotros, pueden proporcionarnos objetivos.[15] Discutamos primero esta apelación a la intuición, y luego pasemos a algunos argumentos a favor de la concepción humeana.

Estoy tan dispuesto como el que más a considerar las intuiciones. Pero no me convence la intuición generalizada que apoya la concepción humeana. Esta intuición es general; es decir, su contenido es una generalización universal "Todas las razones tienen tal y tal forma" o "Ninguna creencia puede ser intrínsecamente motivadora"). Creo que las intuiciones de este tipo a menudo pueden ser fácilmente socavadas por ejemplos particulares de cosas que desafían la generalización. Por ejemplo, consideremos el principio de que "dos objetos sólidos no pueden ocupar el mismo espacio al mismo tiempo". A primera vista, parece evidente. Pero consideremos este caso: tenemos un trozo de arcilla, que podemos llamar "masa". Moldeamos la arcilla hasta formar una estatua. Llamemos a la estatua "Venus". Parece que Lumpy sigue existiendo; simplemente se le ha dado otra forma. Además, Lumpy y Venus ocupan ahora el mismo espacio. Ambos son objetos sólidos. Y son objetos distintos, ya que Lumpy existió antes que Venus. Por lo tanto, tenemos dos objetos sólidos ocupando el mismo espacio al mismo tiempo. ¿No es esto una profunda paradoja filosófica?

No, no lo es. Deberíamos responder al enigma diciendo: "Vale, si vas a contar casos así, entonces dos objetos sólidos pueden ocupar el mismo espacio al mismo tiempo". Cuando dije que esto era imposible, tenía en mente pares de objetos como: mi mano y esta mesa. Si hubiera pensado en "pares de objetos" como la estatua y el vaso de arcilla del que está compuesta, nunca habría dicho que dos objetos sólidos no podían ocupar el mismo espacio, porque obviamente esas cosas sí pueden. Por lo tanto, creo que la intuición general con la que empezamos -aunque fuera una intuición bastante firme y ampliamente compartida- no tiene ninguna fuerza probatoria contra el juicio de que Lumpy y Venus están en el mismo lugar, o que son dos objetos sólidos. La razón es que la intuición general dependía de que tuviéramos en mente ciertos tipos de casos, y no otros, en primer lugar.

No pretendo que esto sea cierto para todas las intuiciones generales, pero sí para muchas de ellas. Y en particular, creo que es cierto de la intuición que apoya la Concepción Humeana. Cuando alguien propone el principio general "las creencias no pueden ser motivadoras por sí mismas", creo que sólo podemos juzgar la plausibilidad de esa generalización recorriendo diferentes tipos de creencias en nuestra mente. Sin duda, el principio parece plausible cuando pensamos en la mayoría de las creencias, por ejemplo, la creencia de que mi casa está en llamas, la creencia de que las comadrejas son mamíferos, la creencia de que 13 es primo, etc. De hecho, parece plausible cuando se considera cualquier creencia sobre cualquier hecho puramente natural y externo. Pero consideremos ahora otra clase de creencias: las creencias sobre el valor. Por ejemplo, la creencia de que robar es moralmente incorrecto, la creencia de que la felicidad es buena, etc. Si contamos con este tipo de creencias, el principio no parece en absoluto obvio; de hecho, parece muy inverosímil. Creo que el principio de que sólo los deseos pueden motivar parece obvio sólo cuando ignoramos ciertos casos, al igual que el principio de que dos objetos sólidos no pueden ocupar el mismo espacio sólo parece obvio cuando ignoramos ciertos casos.

Consideremos por qué parece plausible que una creencia sobre un hecho natural no pueda ser intrínsecamente motivadora. Una de las razones es que no está claro qué determinaría el curso de acción que se supone que favorece el hecho natural. Así, supongamos que usted ve a alguien sentado en una estación de tren. Le informas de que el tren a Nueva Brunswick está a punto de salir. ¿Puede este hecho, por sí mismo, motivarle a actuar de alguna manera, por ejemplo, a levantarse y subir al tren? Una de las razones para negar que pueda hacerlo es que no parece haber ninguna acción concreta que guarde una relación especial con el hecho. Por supuesto, el hecho de que el tren esté a punto de salir implica que es necesario levantarse y subir al tren para llegar a Nueva Brunswick

en ese tren. Por otro lado, abstenerse de levantarse y subir al tren es necesario para no llegar a Nueva Brunswick. El mero hecho de que el tren esté a punto de salir no nos dice si llegar a Nuevo Brunswick es bueno o malo, por lo que no escoge ningún curso de acción concreto como el que hay que tomar.

Pero los hechos evaluativos son obviamente diferentes. Nos dicen lo que es bueno o malo, y no hay ningún rompecabezas sobre qué curso de acción apoya un juicio evaluativo. El juicio de que uno debe cumplir las promesas tiene obviamente una relación especial con el cumplimiento de las promesas que no tiene con otros tipos de acciones.

Consideremos ahora el principal argumento de Hume contra la concepción racionalista: Hume dice que la función de la razón es sólo determinar lo que es verdadero y falso. Pero las acciones no pueden ser verdaderas o falsas. Por lo tanto, las acciones no pueden ser evaluadas por la razón, y por lo tanto no pueden ser racionales o irracionales.[16]

Este es un argumento terrible. Consideremos un argumento paralelo: La razón sólo puede juzgar la verdad y la falsedad. Los números no pueden ser verdaderos o falsos. Por lo tanto, la razón no puede hacer ningún juicio sobre los números -por ejemplo, la razón no puede decirnos que el siete es primo. Es de suponer que Hume diría que esto es una aplicación errónea de su principio. Presumiblemente, diría que aunque los números en sí mismos no pueden ser verdaderos o falsos, las proposiciones sobre ellos, incluyendo la proposición de que el siete es primo, pueden ser verdaderas o falsas; y por eso el juicio de que el siete es primo puede caer bajo el ámbito de la razón. Pero si eso es correcto, entonces se aplica una respuesta paralela al argumento de Hume: por supuesto que las acciones en sí mismas no pueden ser verdaderas o falsas, pero las proposiciones sobre ellas, incluyendo la proposición de que alguna acción en particular es correcta, pueden serlo; y es por eso

que el juicio de que una acción es correcta puede caer bajo el ámbito de la razón.

*7.4.2 El argumento de Smith

Michael Smith proporciona un argumento similar en espíritu al de Hume, pero más fuerte. El argumento de Smith comienza con una noción técnica, la de "dirección de ajuste". He aquí una historia para ilustrar esa noción. Hay dos hombres en un supermercado. Uno de ellos tiene una lista de la compra y va por el supermercado recogiendo los artículos de la lista. El otro es un detective que sigue los movimientos del primero. El detective hace una lista de las cosas que compra el comprador. Cuando el comprador llega a la caja, ambos hombres tienen listas idénticas en sus manos. Además, ambos quieren que los artículos de su lista coincidan con los que el comprador tiene en su cesta. Pero hay una diferencia: el comprador debe hacer que el mundo coincida con su lista; el detective debe hacer que su lista coincida con el mundo. Así, si el detective se da cuenta de que su lista dice "Pastelillos", pero el comprador no tiene pastelillos en el carrito, puede solucionar el problema simplemente tachando la palabra "Pastelillos". Pero si el comprador se da cuenta de que su lista dice "Pastelillos" pero no tiene pastelillos en el carrito, no puede solucionar el problema tachando "Pastelillos", sino que tiene que volver a poner algunos pastelillos en el carrito. Podemos expresar esta diferencia diciendo que la lista del comprador tiene una dirección de ajuste de mundo a lista; la lista del detective tiene una dirección de ajuste de lista a mundo.[17]

Ahora bien, se puede decir algo análogo sobre la diferencia entre creencias y deseos: Las creencias, podríamos decir, tienen una "dirección de ajuste mente-mundo", lo que significa que las creencias pretenden corresponder a la realidad. Si la creencia no se corresponde con la

realidad (es decir, es falsa), entonces la creencia es defectuosa y debe cambiarse. Los deseos, en cambio, tienen una "dirección de ajuste mundo-mente", lo que significa que pretendemos que la realidad se corresponda con nuestros deseos. Si un deseo no se corresponde con la realidad (es decir, no se satisface), el deseo no es por ello defectuoso, sino que uno tiene motivos para intentar cambiar el mundo.[18]

Con estos antecedentes, podemos enunciar el argumento de Smith contra la concepción racionalista de las razones motivadoras:

1. Tener una razón motivadora requiere tener una meta.
2. Tener una meta es estar en un estado con dirección mundo-mente de ajuste.
3. Las creencias tienen una dirección de ajuste mente-mundo, no mundo-mente.
4. Por lo tanto, las creencias no pueden ser razones motivadoras.[19]

(4) entra en conflicto con la concepción racionalista, ya que los racionalistas dicen que una creencia moral puede constituir una razón motivadora.

¿Dónde podría un racionalista cuestionar este argumento? Creo que la premisa crucial es (2).[20] En primer lugar, aclaremos (2) sustituyendo la expresión metafórica "dirección de ajuste del mundo a la mente" por su significado literal. Considero que el significado literal de (2) es más o menos éste:

> Tener una meta es estar en un estado tal que, si uno se da cuenta de que el mundo no coincide con ese estado, entonces tendrá una tendencia a tratar de cambiar el mundo para eliminar la discrepancia.[21]

Pero esto sigue siendo ambiguo. ¿Qué significa que un estado mental

"coincida" o "se ajuste" al mundo? Podría significar que el contenido del estado mental se realiza, o que la meta que el estado mental representa se realiza. Esto da lugar a las siguientes dos lecturas de (2):

2_C. Tener una meta es estar en un estado tal que, si uno nota que el contenido de ese estado no se realiza, entonces uno tendrá una tendencia a tratar de cambiar el mundo para lograr el contenido de ese estado.

2_G. Tener una meta es estar en un estado tal que, si uno nota que la meta representada por ese estado no se realiza, entonces uno tendrá una tendencia a tratar de cambiar el mundo para lograr esa meta.

Para ilustrar la distinción, supongamos que tengo el objetivo de aumentar mi felicidad. Según la concepción racionalista, el logro de este objetivo podría consistir simplemente en creer que mi felicidad es buena.[22] Ahora bien, el contenido de esta creencia -la cosa que creo que es el caso- es: mi felicidad es buena. Pero la meta que el estado mental representa es simplemente mi felicidad. Esto es sólo parte de la visión racionalista -la visión racionalista es que la meta de uno es la cosa que uno cree que es buena. Por lo tanto, (2_C) implica: si me doy cuenta de que la felicidad no es de hecho buena, entonces tendré la tendencia a intentar que la felicidad sea buena. (2_G), por otro lado, implica que, si me doy cuenta de que no soy feliz de hecho, entonces trataré de hacer que sea feliz.

Obviamente, si (2) significa (2_C), entonces el racionalista debe rechazar la premisa. Si de alguna manera descubro que la felicidad no es buena, no trataré de hacer que la felicidad sea buena (¿cómo podría hacerlo? No puedo cambiar los hechos de la ética). Por otro lado, si (2) significa (2_G), deberíamos aceptarla, pero entonces deberíamos rechazar la premisa (3). Pues, para evitar el equívoco, ahora habrá que leer (3):

3_G- Si uno nota que la meta representada por una creencia no se realiza, entonces tendrá la tendencia a modificar su mente, no a tratar de cambiar el mundo para lograr la meta.

Esto es falso. Si me doy cuenta de que mi felicidad (el objetivo representado por mi creencia "mi felicidad es buena") no se realiza, entonces tendré tendencia a intentar cambiar el mundo para conseguir ese objetivo, es decir, para conseguir que yo sea feliz.

¿Por qué Smith no se da cuenta de la ambigüedad de (2)? Tal vez porque, desde su punto de vista de la motivación, no existe tal ambigüedad. Desde el punto de vista humeano, cuando uno tiene una meta, el contenido de su estado mental es lo mismo que la meta que representa. Por ejemplo, cuando quiero apoderarme del mundo, el contenido del deseo es "que me apodere del mundo", y la meta que representa es también que me apodere del mundo. Por tanto, desde el punto de vista humeano, no hay distinción entre (2_C) y (2_G). Pero hay una distinción en el punto de vista racionalista. Cuando uno tiene una creencia moral, si el objetivo que se representa es *G*, entonces el contenido de la creencia no es *G*, sino "*G* es bueno".

Nótese que, en esta sección, no estoy tratando de refutar la Concepción Humeana. Sólo intento mostrar que Smith no logra refutar la Concepción Racionalista. Para que Smith tenga éxito, no puede asumir la Concepción Humeana para empezar. Por lo tanto, no puede asumir que (2_C) y (2_G) son equivalentes. Así que debe decirnos a cuál de ellas se refiere. Si tomamos en serio la concepción racionalista, no hay ninguna razón por la que aceptaríamos tanto (2) como (3) del argumento de Smith: dependiendo de cómo se lea la noción de que un estado mental "concuerda" con la realidad, encontraremos (2) obviamente inverosímil (si se lee como [2_C]) o encontraremos (3) obviamente inverosímil (si se lee como [3_G]). Por lo tanto, Smith no proporciona un argumento contra

el racionalista que no suscite dudas.

Hasta ahora, he criticado las motivaciones de la Concepción Humeana. En las dos secciones siguientes, me propongo demostrar su profunda inverosimilitud.

7.5 ¿Ampliación de la concepción humeana?

El propio David Hume sostenía una versión muy simple de la Concepción Humeana. En su opinión, es imposible que cualquier acción o deseo sea irracional; sólo las creencias pueden ser evaluadas por la razón:

> *Las acciones pueden ser admirables o censurables, pero no pueden ser razonables o irracionales.*[23]

> *No es contrario a la razón preferir la destrucción del mundo entero a rascarme el dedo. No es contrario a la razón que yo prefiera mi ruina total, para evitar el menor malestar de un indio o de una persona totalmente desconocida para mí. Tampoco es contrario a la razón preferir incluso mi propio y reconocido bien menor a mi bien mayor.*[24]

Esto es absurdo. Incluso los humeanos modernos no siguen a Hume tan lejos. Los ejemplos de acciones irracionales son fáciles de encontrar; de hecho, lo llamativo que ha hecho famosa la segunda cita anterior proviene precisamente del hecho de que los tres ejemplos que Hume da de cosas "no contrarias a la razón" se considerarían normalmente paradigmas de irracionalidad. La única justificación que ofrece Hume para sus valoraciones contraintuitivas de esos ejemplos consiste en una deducción de la concepción humeana de las razones. Pero, ¿por qué deberíamos aceptar esa concepción? Si una teoría de la racionalidad

práctica propuesta es incapaz de dar cuenta de ejemplos aparentemente centrales de irracionalidad práctica -y especialmente si es tan anémica como para ser incapaz de trazar cualquier distinción en absoluto entre el comportamiento racional e irracional- entonces es difícil ver por qué deberíamos tomar esa teoría en serio.

Afortunadamente, los humeanos modernos se afanan en asegurarnos que la concepción humeana no implica tales resultados. Por supuesto, existen decisiones razonables y no razonables, y lo que es racional no está determinado simplemente por el deseo más fuerte del agente en el momento de la elección. Más bien, hay un proceso de deliberación racional, cuyo resultado determina cuál es el curso de acción racional.[25] Este punto de vista sigue siendo ampliamente humeano, porque se considera que el proceso de deliberación racional no genera ninguna meta nueva independientemente de los deseos del sujeto. En lo sucesivo, me referiré a cualquier punto de vista que acepte aproximadamente la concepción humeana al tiempo que añade algunas restricciones a la acción racional que van más allá de "elige lo que más quieras ahora", el "punto de vista neohumeano". Consideraré cinco condiciones sobre las razones normativas que han sido propuestas por los neohumeanos:

i) *La restricción de la previsión*: Las razones normativas que tenemos, así como su fuerza, vienen determinadas no sólo por nuestros deseos actuales, sino también por los deseos que sabemos que tendremos en el futuro.[26] Por tanto, es irracional realizar una acción para satisfacer un deseo actual, si se sabe que en el futuro se frustrarán más deseos y más fuertes. Por eso, el estudiante del ejemplo de la sección 7.3 sería irracional si se pasara la noche de fiesta.

ii) *La restricción de la imaginación*: Lo que tenemos como razón normativa para hacer está determinado en parte por lo que desearíamos, si imagináramos situaciones relevantes de forma vívida y precisa.[27] Por ejemplo, supongamos que, si imagináramos

vívidamente la situación de los africanos hambrientos, desearía ayudarles enviando dinero a Oxfam; entonces tengo actualmente una razón no racional para ayudarles.

iii) *La restricción de consistencia*: Si uno tiene un conjunto de deseos incoherentes -es decir, deseos que lógicamente no pueden satisfacerse todos-, entonces tiene una razón no normativa para cambiar sus deseos y hacerlos coherentes. Por ejemplo, si descubro que quiero inyectarme heroína todos los días, pero también quiero no ser un consumidor habitual de drogas, entonces tengo una razón para intentar erradicar uno de esos deseos.

iv) *La restricción de coherencia*: Tenemos una razón normativa para modificar nuestros deseos de forma que se obtenga un conjunto de deseos más coherente, del mismo modo que tenemos razones para modificar nuestras creencias de modo que produzcan un sistema de creencias más coherente, donde la coherencia es una cuestión de lo sistemático y unificado que es un conjunto de deseos/creencias, el grado en que algunos deseos/creencias pueden explicar otros, y cosas por el estilo. Por tanto, es racional adoptar nuevos deseos que ayuden a explicar y sistematizar nuestros deseos existentes, y renunciar a los deseos que no encajan bien con el resto de nuestros deseos.[28] Por ejemplo, supongamos que deseamos el bienestar de todos los seres humanos, excepto de los miembros de una raza específica, y que no tenemos ninguna explicación para descartar a esa raza. Entonces podría conseguir un conjunto de deseos más coherente renunciando a sus prejuicios raciales y pasando a desear el bienestar de todos los seres humanos.

v) *La restricción de la deliberación*: Tenemos una razón normativa para, antes de actuar, emprender una deliberación racional sobre lo que mejor satisfará nuestros deseos, y tenemos una razón normativa para hacer las cosas que esa deliberación consideraría más adecuadas para satisfacer nuestros deseos. Por lo tanto, una persona actúa

> irracionalmente si hace lo que le parece más adecuado para satisfacer sus deseos, cuando una deliberación más profunda habría mostrado que otro curso de acción es mejor.

El punto de vista neohumeano, argumentaré, es inestable. Una vez que adoptamos una concepción ampliamente humeana de las razones, ninguna de estas restricciones de racionalidad puede ser defendida. Un humeano se queda así con la opinión, o bien de que no hay tal cosa como una acción racional o irracional, o bien de que es racional simplemente hacer lo que uno más quiere hacer en ese momento.

7.5.1 La restricción de la previsión

Supongamos que actualmente no tenemos ningún deseo de que *A* ocurra, sino que deseamos que *A* no ocurra; sin embargo, sabemos que en algún momento futuro, nuestro conjunto psicológico cambiará de tal manera que llegaremos a desear *A* (pero no como resultado de ningún cambio en las creencias). ¿Tiene usted ahora una razón -ya sea motivadora o normativa- para tomar medidas para que se produzca *A*? Un humeano debería responder que no. Sólo los deseos de un individuo en el momento de la decisión pueden ayudar a explicar sus acciones; un deseo que tendrá en el futuro pero que no tiene ahora no puede motivar ninguna acción presente, como tampoco puede hacerlo un deseo que solía tener pero que ya no tiene. Si aceptamos los argumentos de la sección 7.2 para una fuerte conexión entre las razones motivadoras y las normativas, esto sugiere que tal deseo futuro tampoco le da razones normativas, como tampoco lo hace un deseo meramente pasado. Compárese un argumento de Bernard Williams: Williams sostiene que para que algo sea una razón normativa para que *S* haga *A*, debe ser algo que pueda ser la razón real de *S* para hacer *A*; pero sólo algo que incluya uno o más de los deseos de *S* podría ser la razón de *S* para hacer

cualquier cosa; por lo tanto, las razones normativas deben implicar deseos.[29] *A* esto podríamos añadir que sólo algo que incluya uno o más de los deseos presentes de *S* podría ser la razón de *S* para hacer algo ahora; por lo tanto, el mismo argumento lleva a la conclusión de que las razones normativas deben implicar deseos presentes.

La historia de Odiseo y las sirenas ilustra este punto: Odiseo planea pasar por delante de las sirenas, cuyos cantos han atraído a la muerte a muchos marineros. Ahora desea firmemente que no le ocurra lo mismo. Pero Odiseo sabe que una vez que escuche el canto de las sirenas, le invadirá el deseo de navegar hacia ellas, pase lo que pase. Este deseo, supongamos, será tan fuerte que considerará que la muerte es un precio que vale la pena pagar para estar más cerca de las sirenas. Pero, repito, ahora desea no ir hacia las Sirenas. Su solución: se hace atar al mástil mientras navega junto a las sirenas. ¿Es Odiseo racional al hacer esto? ¿El deseo que tendrá cuando escuche el canto de las sirenas le proporciona ahora una razón normativa para tomar medidas que le aseguren que podrá unirse a las sirenas? Considero que las respuestas son claramente "sí" y "no". Esto encaja con lo que he dicho que debe decir el humeano.

He aquí otra ilustración del mismo punto. Sunflower es actualmente un estudiante universitario socialista. Para su consternación, ha observado que muchos antiguos socialistas se han "vendido" al hacerse mayores, cambiando sus valores para ajustarse a los de su sociedad. Supongamos que Sunflower cree justificadamente que, cuando se haga mayor, su propia personalidad y sus valores también cambiarán de forma que la llevarán, sin adquirir ninguna falsa creencia no valorativa, a respaldar el capitalismo. Pero, repito, actualmente se opone al capitalismo y apoya el socialismo. ¿Tiene ahora Girasol una razón para trabajar por el mantenimiento del orden capitalista? Supongo que la

respuesta correcta, y la que el humeano aceptará, es "no".

A continuación, supongamos que usted tiene actualmente un deseo de cierta fuerza de que ocurra *A*, pero sabe que en un momento futuro su deseo será más fuerte o más débil que el actual. ¿Cuántas razones tiene ahora para intentar que se produzca *A*? Es obvio que la fuerza con la que estoy motivado para hacer algo debe estar determinada por la fuerza de mis deseos actuales, no por la fuerza que tengan o tengan en otro momento. Y dados los argumentos de la sección 7.2, esto sugiere que lo mismo es cierto para nuestras razones normativas. Además, esta respuesta parece requerida por lo que ya hemos dicho: una vez que aceptamos que lo que ahora tenemos razones para hacer está determinado por lo que actualmente deseamos, seguramente deberíamos decir también que cuántas razones tenemos ahora para hacer una cosa están determinadas por la fuerza actual de nuestros deseos.

Pero ahora hemos renunciado por completo a la restricción de previsión. Sólo los deseos presentes son relevantes, tanto para las razones que tenemos como para la fuerza de esas razones. Los deseos futuros son irrelevantes. O mejor dicho, los deseos futuros sólo pueden ser relevantes de la misma manera que cualquier otro hecho del mundo. Si tengo un deseo actual de que se satisfagan mis deseos futuros, entonces el hecho de que en el futuro desee *x* es relevante para mi acción, pero sólo del mismo modo que si deseo que se satisfagan los deseos de Zbigniew Brzezinski, entonces el hecho de que Zbigniew Brzezinski desee *x* es relevante para mi acción. En ambos casos, es mi deseo presente el que proporciona la motivación para mi acción, y en ambos casos, podría haber carecido de ese deseo y seguir siendo perfectamente racional. Del mismo modo que los humeanos dicen que no hay ninguna razón racional por la que yo deba desear que se satisfagan los deseos de otra persona, también deberían decir que no hay ninguna razón racional

por la que yo deba desear que se satisfagan mis propios deseos futuros.[30]

Por lo tanto, no hay ninguna razón aparente por la que el estudiante de mi ejemplo anterior sea irracional para pasar la noche de fiesta. Si decide hacerlo, será porque su deseo de salir de fiesta era, en el momento en que tomó la decisión, más fuerte que su deseo de aprobar la clase, graduarse, etc. (¿cómo podría entenderse su acción si no?). Aunque sabía que más tarde su deseo de aprobar la asignatura sería mucho más fuerte, esto es irrelevante. Incluso si la estudiante tuviera el deseo de que sus deseos futuros fueran satisfechos, esto no haría que su acción fuera irracional; el hecho de que eligiera salir de fiesta mostraría simplemente que, en el momento en que tomó la decisión, su deseo de salir de fiesta era más fuerte.

7.5.2 La restricción de la imaginación

La restricción de la imaginación sostiene que nuestras razones para actuar están en parte en función de los deseos que no tenemos pero que tendríamos si imagináramos vívidamente y con exactitud determinadas situaciones. La imaginación vívida puede alterar nuestros deseos haciéndonos notar características de una situación que de otro modo pasaríamos por alto, o permitiéndonos comprender mejor una situación que de otro modo no entenderíamos; entonces podríamos tener nuevas reacciones conativas ante la situación. De estas y otras maneras, la imaginación vívida podría hacernos experimentar diferentes emociones o apetitos, por ejemplo, despertando en nosotros un sentimiento de simpatía (no nos ocupamos aquí de los efectos puramente cognitivos de la imaginación).

Sin embargo, no está claro por qué un deseo que no tengo pero que tendría si realizara un ejercicio imaginativo debería considerarse relevante para mis razones prácticas reales. Un deseo meramente

potencial no puede servir como razón motivadora; sólo los deseos reales pueden moverme a actuar. Dada la conexión entre las razones motivadoras y las normativas, parece que tampoco puede servir como razón normativa. Es cierto que imaginar una situación de forma vívida puede cambiar mi estado emocional y, por tanto, mis deseos, pero ¿qué hace que el nuevo conjunto de deseos sea mejor o "más racional" que el anterior? Tal vez si la corrección de algunos errores en mi razonamiento modificara mis deseos, habría un sentido inteligible en el que los nuevos deseos serían más racionales, serían más un producto del uso correcto de la razón. Pero este no es el tipo de cambio que estamos discutiendo aquí.

Supongamos un ejemplo: mi médico me ha dicho que tengo que operarme para extirparme el apéndice. Acabo de llegar al hospital el día previsto para la operación. Sé que, debido a mi carácter aprensivo, si me imaginara el procedimiento quirúrgico con todo detalle, me horrorizaría tanto la imagen del cirujano abriendo el apéndice, etc., que sentiría un impulso irrefrenable de huir del hospital. Sin embargo, en la actualidad no tengo ese impulso y deseo seguir adelante con la operación. ¿Qué es lo racional para mí? Parece claro que debo seguir adelante con la operación, y sólo abstenerme de imaginarla vívidamente. En este caso, el deseo que tendría si imaginara vívidamente la operación no me da ninguna razón para actuar.

Pero volvamos al ejemplo de Oxfam y los africanos hambrientos. Supongamos que no quiero dar dinero a Oxfam, porque disfrutaría mucho más gastando el dinero en un nuevo equipo de música. Pero también sé que si imaginara vívidamente la situación en África, me sentiría tan mal por ello que acabaría enviando mucho dinero a Oxfam, en lugar de disfrutar del nuevo equipo de música. En una concepción húngara de las razones, ¿por qué no debería concluir simplemente: "Bueno, entonces será mejor que no imagine vívidamente la situación en

África", y seguir adelante con mis planes de compra del equipo de música?

7.5.3 Las restricciones de consistencia y coherencia

Al menos el humeano puede defender un requisito racional de consistencia y coherencia para nuestros deseos, del mismo modo que la razón requiere creencias consistentes y coherentes, ¿no es así?

No, no puede. Primero, revisemos la diferencia básica entre creencias y deseos. Recordemos que Michael Smith, el mismo filósofo que quiere defender un requisito de coherencia para los deseos, caracteriza la diferencia entre creencias y deseos en términos de "dirección de ajuste". Las creencias tienen una "dirección de ajuste mente-mundo", lo que significa que pretenden ajustarse a la realidad. Los deseos tienen una "dirección de ajuste mundo-mente", lo que significa que pretendemos que la realidad se ajuste a nuestros deseos.

En segundo lugar, repasemos qué tiene de malo la incoherencia en las creencias. Una creencia incoherente es aquella que no puede ser verdadera. Puesto que las creencias tienen una dirección de ajuste mente-mundo, esto las señala como intrínsecamente defectuosas; siempre hay una razón para abandonar tales creencias. Si un (pequeño) conjunto de creencias es incoherente, uno debería cambiar sus creencias para eliminar la incoherencia, por la misma razón, es decir, para alcanzar un sistema de creencias más exacto.[31] Pero esta línea de pensamiento no se aplica a los deseos. Si un deseo es incoherente, no puede satisfacerse. Pero como los deseos tienen una dirección de ajuste mundo-mente, la insatisfacción de un deseo no lo marca como defectuoso. Puesto que, es decir, intentamos (o tenemos motivos para intentar) que la realidad se ajuste a nuestros deseos en lugar de que nuestros deseos se ajusten a la realidad, el hecho de que un deseo no se ajuste, o incluso no pueda

ajustarse, a la realidad no es motivo para cambiar el deseo. Por tanto, la imposibilidad de satisfacer todos los miembros de un conjunto de deseos tampoco hace que ese conjunto de deseos sea defectuoso.

Para ilustrarlo, supongamos que yo creo que Dios existe y quiero que exista. Pero un ateo me da un argumento que demuestra que Dios no puede existir.[32] En este punto, se demuestra que mi creencia ha sido gravemente defectuosa (es decir, errónea), y si soy racional, debo dejar de tener esa creencia. Pero con ello no se demuestra que mi deseo haya sido "erróneo" en ningún sentido, ni soy irracional si sigo deseando que haya un Dios, ni hay nada defectuoso en un sistema de valores según el cual sería bueno que hubiera un Dios.

Debemos distinguir este punto de algunas afirmaciones superficialmente similares pero falsas. Como he dicho, la insatisfacción de un deseo no constituye ningún defecto del deseo. Sin embargo, no niego que la imposibilidad de llevar a cabo una intención, realizar una esperanza o tener éxito en un intento constituya una razón racional para renunciar a la intención, esperanza o intento. ¿Por qué? Porque las intenciones, las esperanzas y los intentos implican algo más que deseos. La intención de hacer *A* implica la creencia de que se hará *A* (por ejemplo, "tengo la intención de ganar esta carrera" frente a "quiero ganar esta carrera"). La esperanza de que p implica la creencia de que p es posible. Y un intento, además de implicar quizás el mismo tipo de creencia que una intención, normalmente requiere, como parte de su justificación práctica, la creencia de que el intento ayudará o podría ayudar a conseguir algún objetivo; pero los intentos fallidos no suelen ayudar a conseguir los objetivos. Por eso las intenciones, las esperanzas y los intentos son criticables desde el punto de vista de la imposibilidad de sus objetos, mientras que los deseos no lo son.

Sin embargo, uno podría pensar que al menos algunos deseos son

irracionales porque sus objetos son imposibles. Por ejemplo, parece que sería irracional querer que 2 sea igual a 3, o desear que llueva y no llueva al mismo tiempo. Creo que si estos deseos son irracionales, no es sólo porque no puedan satisfacerse; tal vez sean irracionales porque es ininteligible cómo sus objetos podrían ser buenos, o porque uno no puede concebir coherentemente sus supuestos objetos, lo que hace imposible que un ser plenamente racional tenga esos deseos. Pero en cualquier caso, estos casos apenas nos conciernen, ya que tales deseos ininteligibles son extremadamente raros, si es que existen. La cuestión en esta sección es si el humeano puede proporcionar suficientes restricciones para darnos una explicación plausible de la acción racional. La capacidad de criticar estos deseos ininteligibles seguiría dejando al humeano con una concepción muy limitada de las razones normativas, lo que nos llevaría a muy poca distancia hacia una noción robusta de racionalidad práctica.

Debo abordar un argumento a favor de la afirmación de que tenemos una razón normativa para eliminar los deseos incoherentes: Queremos que nuestros deseos, en general, estén bien satisfechos, ya que cuantos más de nuestros deseos estén satisfechos, más felices seremos, y queremos ser felices. Si tenemos deseos incoherentes, sabemos que no serán satisfechos. Por lo tanto, tenemos una buena razón para intentar cambiar nuestros deseos por otros que sean coherentes. Se trata de un razonamiento directo de medios-fines: debemos eliminar los deseos incoherentes, porque al hacerlo contribuiremos a alcanzar un fin que deseamos. Obsérvese que, si este argumento fuera correcto, se aplicaría no sólo a los deseos que no pueden satisfacerse, sino también a los deseos que tienen pocas probabilidades de satisfacerse o que son difíciles de satisfacer. Demostraría que tenemos una razón normativa para convertirnos en un conjunto de deseos fácilmente satisfactibles, como el deseo de que los cuadrados tengan cuatro lados, que las rocas

estén hechas de átomos, etcétera. Dado que es poco probable que el hambre en el mundo desaparezca en un futuro previsible, tenemos una razón para erradicar, si es posible, el deseo de que la gente no se muera de hambre. Por supuesto, puede que nos resulte difícil o imposible erradicar ese deseo, pero lo mismo puede ocurrir con cualquier otro deseo. Para el humeano, no hay ninguna diferencia de principio entre los deseos inconsistentes y los deseos difíciles de satisfacer.

Estos últimos resultados son bastante inverosímiles; no pensamos en el deseo de acabar con el hambre en el mundo como algo irracional, y no pensamos que un ser que deseara sólo lo inevitable sería por ello más racional que nosotros.[33] Hay una explicación directa, no humeana, para esto: dado que muchos estados de cosas evitables son buenos, vemos los deseos por ellos como apropiados. Del mismo modo, vemos la indiferencia hacia los estados de cosas malos como algo inapropiado. Por lo tanto, pensamos que la naturaleza del hambre en el mundo es tal que exige una reacción emocional totalmente negativa, y que una persona que carezca de tal reacción sería, por lo tanto, deficiente.

Un humeano podría intentar acomodar esta idea tratándola como una cuestión de deseos adicionales que tenemos -en particular, deseos de que mantengamos nuestros deseos actuales-. Pero si esto funciona, entonces lo mismo podría ocurrir con un agente con deseos inconsistentes: él también puede desear mantener sus deseos actuales. De nuevo, el humeano no tiene ninguna razón teórica para tratar los deseos inconsistentes de forma diferente a los deseos meramente difíciles de satisfacer.

Una vez que vemos esto, la idea de Michael Smith de que debemos buscar un conjunto coherente de deseos pierde toda plausibilidad. La razón por la que los epistemólogos consideran deseable un conjunto coherente de creencias es que dicho conjunto tiene más probabilidades

de ser verdadero en general.[34] Pero como la tarea de los deseos no es la de confontar con la realidad, no se puede sostener el argumento paralelo de que hay que preferir un conjunto sistemático y unificado de deseos por el hecho de que es más probable que se satisfaga en general. El hecho de que los deseos tengan una dirección de ajuste entre el mundo y la mente frustra el paralelismo que Smith quiere establecer.

7.5.4 La restricción de la deliberación

Acabo de atacar algunos de los principales principios de la teoría neohumeana de la racionalidad práctica. Lo que queda es tal vez el principio más central y supuestamente inexpugnable de este punto de vista: que al menos tenemos una razón normativa para averiguar qué satisfaría nuestros deseos actuales y adherirnos a los resultados de dicho razonamiento. Pero argumentaré que ni siquiera esto puede servir como restricción que permita criticar racionalmente algunos comportamientos.

Si aceptamos que sólo los deseos actuales y reales pueden figurar en las razones (motivadoras o normativas) para la acción, entonces es plausible que digamos lo mismo sobre las creencias: sólo las creencias actuales y reales de una persona -no las creencias que tendrá en algún momento futuro, ni las creencias que tendría en alguna situación contrafactual- pueden figurar en las razones para la acción que tiene actualmente. Los argumentos parecen ser los mismos: una creencia que uno no tiene ahora no puede tener ninguna influencia en su comportamiento actual. Además, dado esto, parece que esa (no) creencia no puede figurar en ninguna de las razones normativas que uno tiene actualmente.

Para ilustrarlo, supongamos que usted está considerando si va a iniciar una nueva carrera como trapecista en el circo, o si va a seguir con su trabajo como profesor. Tiene mucho tiempo para decidir. Después de

pensarlo durante cinco minutos, le parece que ser trapecista es lo que mejor satisface sus deseos, y decide hacerlo. Pero suponga que, si hubiera deliberado un poco más y con más detenimiento, habría llegado racional y correctamente a la conclusión de que seguir siendo profesor satisfaría mejor sus deseos. ¿Se le puede criticar racionalmente, desde un punto de vista humeano, por unirse al circo? Dividamos esto en dos preguntas:

En primer lugar, dadas sus creencias y deseos en el momento de la decisión, ¿puede ser criticado racionalmente por unirse al circo? No. Usted decidió unirse al circo porque tenía una serie de deseos y pensó que esa decisión sería la que mejor los satisfaría. Si no fuera así -si hubieras esperado que la otra opción satisfaciera mejor tus deseos-, simplemente no habría explicación, dada una concepción Humeana de la motivación, de cómo pudiste acabar eligiendo el circo. La propia explicabilidad de tu acción garantiza su racionalidad en términos de tus creencias y deseos en ese momento. La única manera de que un humeano critique racionalmente su acción, entonces, es criticar sus creencias en el momento de la decisión.

En segundo lugar, entonces, ¿se le puede criticar racionalmente por creer como creía? El humeano quiere decir "sí". Considere dos formas que podría adoptar la crítica: (a) Podría ser criticado por algo sobre lo que no tenía ningún control, por ejemplo, por no ser más inteligente o por no tener alguna información que no tenía forma razonable de obtener. Pero esto no parece aceptable. El tipo de crítica que nos interesa en primer lugar es una acusación de la racionalidad práctica de su decisión de convertirse en trapecista, es decir, una crítica en el sentido de que racionalmente debería haber hecho otra cosa. Una crítica de algo sobre lo que usted no tenía ningún control, independientemente de que esa cosa condujera o no a su acción, no parece relevante para demostrar que

usted debería haber actuado de forma diferente a como lo hizo. (b) Se le puede criticar por algo sobre lo que tenía control, por ejemplo, por no dedicar más tiempo a deliberar o por no esforzarse más en ser cuidadoso. Pero esta crítica fracasará por las razones expuestas en el párrafo anterior. El acto de dedicar cierto tiempo a deliberar, por ejemplo, tendrá que ser evaluado como racional dadas las creencias y deseos que tenías en el momento de realizarlo, si es que la acción es explicable. Si hubieras pensado que deliberar durante más tiempo serviría mejor a tus fines, entonces (dada la explicación humeana de la motivación) lo habrías hecho. Si no pensabas que deliberar más tiempo serviría mejor a tus fines, entonces era racional, dado eso, que no deliberaras más tiempo. Así, el humeano puede criticarte por no deliberar más tiempo sólo si puede criticarte por alguna(s) creencia(s) adicional(es). Se avecina una regresión, en la que en cada etapa el intento de crítica del humeano exige alguna otra creencia por la que se le pueda criticar. En resumen, para cualquier persona, *S*, y cualquier suceso o estado de cosas, *x*:

1. Si *S* no controla *x*, entonces *S* no puede ser criticado racionalmente por *x*.
2. Si *S* controla *x*, entonces *S* puede ser criticado racionalmente por *x*, por motivos humeanos, sólo si hay algo más por lo que *S* pueda ser criticado racionalmente.
3. Por lo tanto, *S* puede ser criticado racionalmente por algo, por motivos humeanos, sólo si hay una serie infinita de cosas por las que *S* puede ser criticado racionalmente
4. Nunca existe tal serie infinita.
5. Así que ninguna persona puede ser criticada racionalmente, por motivos humeanos, por nada.

La premisa (1) es cierta porque no se puede considerar irracional a una

persona por no hacer algo que no podía hacer. (2) es cierta porque, desde el punto de vista humeano, una acción o elección siempre tiene sentido dadas las creencias y deseos del agente; sólo puede ser irracional si una de esas creencias o deseos es irracional. El paso (3) se sigue (suponiendo que no permitamos el razonamiento circular en las críticas a las acciones), porque el "algo más" mencionado en (2) debe ser (por (1)) otra cosa sobre la que *S* tiene control. Pero para que *S* pueda ser criticado por esta segunda cosa, tendrá que haber una tercera cosa sobre la que *S* tenga control y por la que *S* pueda ser criticado racionalmente... y así sucesivamente. Tomo (4) como una premisa intuitivamente obvia. (5) se deduce de (3) y (4).

Se podría objetar que (1) o (2) son falsas, porque se puede criticar racionalmente a una persona por ciertas creencias (digamos, creencias para las que no hay pruebas), sin que haya nada más por lo que se le pueda criticar. Pero, ¿qué tipo de crítica se puede hacer a las creencias? Un tipo de crítica a una creencia sería que la creencia no es cierta, o que no es probable que lo sea.[35] Otro tipo sería una crítica práctica: que uno no debería haber adoptado la creencia, que uno tenía buenas razones normativas para no adoptarla. Pero el primer tipo de crítica parece irrelevante para nuestros fines. Si una persona ha hecho todo lo que tenía sentido que hiciera dadas sus creencias, y no podría haber hecho nada para obtener creencias mejores, entonces no veo cómo una crítica de sus creencias podría mostrar que, a pesar de todo, racionalmente debería haber actuado de otra manera. Consideremos ahora el segundo tipo de crítica. Se podría sostener (aunque la mayoría de los filósofos lo dudan) que podemos controlar nuestras creencias y que tenemos razones prácticas normativas para adoptar creencias epistémicamente racionales.[36] Pero esto nos devolvería a la regresión antes mencionada: tu acto de adoptar una creencia epistémicamente irracional estaría en sí mismo sujeto a la explicación humeana de la racionalidad práctica, en

cuyo caso cualquier crítica de este acto requeriría que hubiera algo más por lo que fueras criticable.

7.5.5 Coda: la necesidad de hechos evaluativos

He argumentado que ninguna de las restricciones de racionalidad propuestas puede motivarse dentro de una concepción ampliamente humeana. Además, en cada caso podemos aducir ejemplos claros, tanto de acciones o deseos racionales que violan la restricción, como de otros irracionales. Por lo tanto, en una teoría de la racionalidad práctica se necesita alguna estructura adicional no contenida en la concepción humeana, para dar cabida a ejemplos paradigmáticos de acciones racionales e irracionales. Esa estructura adicional, sostengo, la proporcionan los hechos evaluativos.

Consideremos de nuevo la restricción de previsión. Odiseo es racional al violar la restricción cuando no da importancia a su deseo futuro de navegar más cerca de las sirenas. Pero el estudiante fiestero es irracional al violar la restricción, si no le da la importancia adecuada a su deseo futuro de aprobar su examen. Para un humeano, esto es paradójico: a veces consideramos que los deseos futuros son relevantes para lo que es racional, y otras veces no. Si los deseos básicos no pueden ser evaluados, entonces parece que deberíamos dejar entrar todos los deseos futuros básicos, o excluir todos los deseos futuros básicos de nuestras razones para la acción. Desde el punto de vista del sentido común, la solución del rompecabezas es obvia: Odiseo juzga racionalmente que su deseo futuro es inapropiado, un deseo de lo que en realidad es malo. Pero el estudiante fiestero juzga racionalmente, o debería juzgar, que sus deseos futuros son apropiados, deseos de lo que en realidad es bueno. Por eso es racional que Odiseo frustre intencionadamente su deseo futuro, pero es irracional que el estudiante haga lo mismo. Ahora bien, el

humeano podría querer argumentar que el deseo anticipado de Odiseo sería erróneo sobre la base de alguna otra restricción de racionalidad neohumeana. Pero este movimiento no está disponible si mis argumentos con respecto a las otras restricciones propuestas han tenido éxito.

En segundo lugar, consideremos la restricción de la imaginación. En el caso de la cirugía, violo racionalmente esta restricción al no conceder ninguna importancia al deseo que formaría si imaginara vívidamente mi próxima cirugía. Pero en el caso de Oxfam, parece que violo irracionalmente la restricción cuando no doy importancia al deseo que formaría si imaginara vívidamente la situación de los africanos hambrientos. De nuevo, esto es un rompecabezas para los humeanos, pero la explicación del sentido común es sencilla: el deseo contrafactual en el caso de la cirugía sería un deseo no guiado, un deseo por lo que es malo. El deseo contrafactual en el caso de Oxfam sería un deseo adecuado, un deseo por lo que es bueno. Por eso nos parece racional eludir el primer deseo absteniéndose de pensar demasiado en la cirugía, pero no nos parece racional eludir el segundo deseo absteniéndose de pensar demasiado en el hambre en el mundo.

En tercer lugar, consideremos la restricción de consistencia. En el supuesto de que Dios no puede existir, mi deseo de que Dios exista violaría racionalmente esta restricción o alguna otra estrechamente relacionada. Por otro lado, el deseo de que dos sean tres parece violar irracionalmente la restricción. ¿Por qué? Porque incluso si suponemos que la existencia de Dios es imposible, podemos entender cómo la existencia de un ser perfecto sería buena. Pero no podemos entender cómo dos es igual a tres sería bueno.

Por último, recordemos la restricción de coherencia de Michael Smith. Si mis preferencias estéticas son eclécticas, no por ello soy irracional en ninguna medida. Por otra parte, se ve que el racista que se

preocupa por todos los seres humanos excepto los miembros de una raza específica es por ello irracional. Los hechos evaluativos ayudan de nuevo a explicar esto: mis gustos estéticos no tienen por qué reflejar ningún hecho moral. Si me gusta una pieza musical pero me disgusta otra muy similar, no tengo por qué sostener la opinión de que una pieza tiene una propiedad evaluativa objetiva (o incluso estética) de la que la otra carece. Sin embargo, pensamos que la consideración del bienestar de las personas debería reflejar, al menos en parte, el valor de los seres humanos. Dado que no podemos encontrar ninguna distinción moralmente relevante entre una raza y otra, pensamos que las personas de todas las razas tienen el mismo valor como seres humanos. Por lo tanto, la falta de consideración hacia los miembros de una raza demuestra una ceguera ante ese valor. Los casos en los que la restricción de coherencia parece apropiada pueden entenderse como casos en los que pensamos que hay hechos evaluativos que limitan los deseos apropiados.

Al evaluar la promesa de este enfoque, recomiendo al lector dos preguntas. En primer lugar, ¿son plausibles las valoraciones de racionalidad práctica que he hecho sobre los ejemplos anteriores? En segundo lugar, ¿es plausible que nuestra tendencia a aceptar esas evaluaciones se explique por nuestra aceptación de juicios evaluativos sobre la idoneidad de los deseos, del tipo que he sugerido? Si es así, entonces la visión humeana estándar de la relación entre los deseos y las evaluaciones es un error. El punto de vista humeano es que es racional actuar de acuerdo con nuestros juicios evaluativos sólo porque hacerlo ayuda a satisfacer nuestros deseos. Yo sostengo que es racional actuar de acuerdo con nuestros deseos sólo cuando estamos justificados al juzgar que la satisfacción de esos deseos es buena. Los juicios valorativos dan fuerza normativa a los deseos, no al revés.

Esta conclusión nos orienta claramente hacia el intuicionismo ético. Porque si hay hechos que proporcionan razones normativas para la acción, parece que deben pertenecer a una de estas tres categorías: (a) hechos sobre nuestra psicología; (b) hechos sobre el mundo externo, natural; o (c) alguna especie de hechos objetivos no naturales. Mis críticas a la concepción humeana muestran que (a) es insuficiente para explicar todas nuestras razones normativas, dado que los únicos hechos psicológicos que se consideran plausibles para proporcionar razones para la acción son los hechos sobre los deseos. Además, en general se acepta que los hechos del tipo (b) no proporcionan por sí mismos razones para la acción (véase la sección 7.4). Por tanto, sólo queda la (c): algunas de nuestras razones para la acción consisten en hechos objetivos no naturales, como el hecho de que ciertas cosas sean buenas o malas. Pero sólo los intuicionistas pueden aceptar esta respuesta, porque sólo los intuicionistas son dualistas metaéticos (véase la sección 1.4): sólo los intuicionistas creen que existen hechos del tipo (c). Cualesquiera que sean sus diferencias, los partidarios de los "otros cuatro puntos de vista metaéticos están de acuerdo en que los únicos hechos que existen son los hechos sobre nuestra psicología y sobre el mundo natural.

T. M. Scanlon parece ofrecer una alternativa a los hechos no naturales de los intuicionistas. Sostiene que podemos hablar de que los juicios evaluativos son verdaderos o falsos sin plantear un reino de hechos evaluativos independientes de nosotros. Todo lo que se necesita es que haya normas para las formas correctas de llegar a los juicios evaluativos. Los juicios evaluativos a los que se llega de acuerdo con las normas son verdaderos; los que se obtienen de forma contraria a las normas son falsos. Por supuesto, una cláusula evaluativa no se convierte en verdadera por su conformidad con cualquier norma arbitraria; debe corresponder a normas correctas de juicio evaluativo. Como reconoce Scanlon, el juicio de que una norma para formar juicios evaluativos es

correcta es en sí mismo un juicio evaluativo del tipo al que se aplica la norma.[37]

Se ve que hay algún tipo de circularidad en el relato de Scanlon, y de hecho creo que esto impide que su punto de vista constituya una alternativa real al intuicionismo. Consideremos un juicio evaluativo:

A. La compasión es buena.

Supongamos que lo que hace que (A) sea verdadera es el hecho de que se puede llegar al juicio (A) de una manera que se ajusta a las normas correctas para llegar a los juicios evaluativos. Abreviaré este hecho como sigue:

B. El juicio (A) se ajusta a las normas correctas.

Ahora bien, el juicio expresado por (B) también es evaluativo. ¿Qué hace que sea verdadero? Para ser coherentes, deberíamos decir:

C. El juicio (B) se ajusta a las normas correctas.

Esto también es evaluativo, y podemos hacer la misma pregunta sobre ello, y así *ad infinitum*. El punto de vista de Scanlon implica, por tanto, que las condiciones de verdad de cualquier enunciado evaluativo son infinitamente complejas.[38] Esto me parece increíble. Además, en ningún punto de la serie infinita de condiciones se elimina el elemento evaluativo; por tanto, el argumento no muestra que se puedan acomodar las verdades evaluativas apelando únicamente a los hechos naturales.

7.6 La autoridad de la moral

7.6.1 La autoridad de la moral: un punto de vista racionalista

Algunas de nuestras razones para actuar se derivan de los deseos,

concretamente de los apetitos y las emociones. Otras, las razones prudenciales e imparciales, se derivan de la cognición de los hechos: el hecho de que una acción sirva mejor a mis intereses a largo plazo constituye una razón prudencial para que tome ese curso de acción; el hecho de que una acción promueva el bien constituye una razón moral imparcial para que tome ese curso de acción.

¿Qué hay que hacer cuando entran en conflicto razones de distinta índole? Parece una pregunta razonable e importante, pero ¿qué significa? Si "¿Qué debe uno hacer?" significa "¿Qué está uno moralmente obligado a hacer?", entonces la respuesta es trivial: obviamente, uno está moralmente obligado a actuar según sus razones morales, si las hay. Cuando formulamos la pregunta, con la intención de que no sea trivial, creo que nuestra concepción es la siguiente: hay diferentes tipos de razones que tenemos para actuar, incluyendo razones emocionales, razones prudenciales y razones morales. No se trata de "razones" en diferentes sentidos de la palabra, sino simplemente de diferentes razones que podemos tener, en algún sentido genérico de "razón". Por lo tanto, estas razones pueden sopesarse entre sí. Así que la pregunta que surge es, cuando sopesamos todas las razones que tenemos, ¿qué es lo que tenemos más razones, en general, para hacer? Por ejemplo, ¿tenemos siempre más razones para ser morales, o a veces es más razonable ser inmoral? Mi opinión es que la prudencia supera a los deseos, y la moral a la prudencia. Por tanto, siempre es más racional actuar moralmente. Esta característica de la moral es lo que yo llamo "la autoridad de la moral". El término sugiere una analogía con la autoridad legal: por ejemplo, en el sistema político estadounidense, la Constitución tiene prioridad sobre los estatutos, que a su vez tienen prioridad sobre los reglamentos; la Constitución tiene la máxima autoridad. Se trata de una afirmación normativa, no descriptiva: no se trata de que, de hecho, nunca se apliquen estatutos que contravengan la Constitución, sino de que tales

estatutos son ilegítimos; se exige legalmente que tales estatutos no se elaboren ni se apliquen. De forma análoga, la moral tiene prioridad sobre la prudencia, que tiene prioridad sobre los deseos. Esto no significa que, de hecho, nunca se lleven a cabo acciones inmorales de forma intencionada. Sólo significa que tales acciones son irracionales; uno siempre tiene razones suficientes para no realizarlas.

¿Por qué la prudencia prevalece sobre los deseos? Porque un juicio prudencial, si se forma correctamente, ya tiene en cuenta los deseos presentes. Este juicio no es en sí mismo un deseo, sino que se hace a la luz de una consideración de lo que uno desea actualmente y de lo fuertes que son esos deseos. Si el juicio prudencial de uno es que debe actuar de tal manera que no satisfaga algún deseo presente, entonces uno ya ha evaluado cualquier otro fin como más importante que la satisfacción de ese deseo. Proceder entonces a sopesar el deseo presente frente a ese juicio sería una doble valoración injustificada del deseo. Nótese la asimetría: los juicios prudenciales implican evaluaciones de los deseos; los deseos no implican evaluaciones de los juicios prudenciales.

No he dado aquí una justificación general de las exigencias de la prudencia, no he intentado explicar por qué una persona racional debe asignar cierto peso a su bienestar futuro. Simplemente lo asumo. Más bien, he explicado por qué, dada la capacidad de juicio prudencial de una persona, no debe concebirse que las consideraciones prudenciales se sopesen con las emociones o los apetitos.

Ahora bien, ¿por qué la moral tiene prioridad sobre el interés propio? La respuesta, sugiero, es estructuralmente la misma: porque el juicio moral, si se forma correctamente, ya tiene en cuenta el interés propio. Esto depende de una tesis ética sustantiva: que la moral no es insensible a los propios intereses de los agentes. Estamos moralmente autorizados a perseguir nuestros propios intereses, dentro de ciertos límites. Por

ejemplo, si uno pasa por delante de una víctima de un accidente que necesita atención médica, está obligado a llamar a una ambulancia, dado que el coste para uno mismo es trivial. Pero si eres testigo de un atraco, no estás obligado a ayudar a la víctima luchando contra el atracador, ya que si lo hicieras te pondrías en grave peligro.

Por lo tanto, si uno juzga que está obligado a actuar en contra de sus propios intereses en alguna ocasión, eso significa que ya ha sopesado su interés particular en las circunstancias frente a cualquier otra consideración que esté en juego, y ha encontrado que su propio interés se ve superado por esas otras consideraciones. Proceder entonces a sopesar el interés propio frente a ese juicio moral sería un doble cómputo injustificado del propio interés. De nuevo, obsérvese la asimetría: los juicios morales implican evaluaciones de los deseos y de los intereses propios; los deseos y los juicios prudenciales no implican evaluaciones de los intereses morales.

No he dado aquí una justificación general de la moral, ni he refutado el egoísmo ético. Simplemente he explicado por qué, dado que uno tiene la capacidad de juicio moral, uno no debería concebir sus juicios morales como si se sopesaran con sus motivos egoístas.

La cuestión aquí no es que las consideraciones morales sean una categoría más amplia de la que las consideraciones prudenciales sean un subconjunto, ni que las consideraciones prudenciales sean una categoría más amplia de la que los deseos sean un subconjunto. Los deseos no son creencias prudenciales; son de distinto tipo. La cuestión es que las creencias prudenciales incluyen juicios intelectuales sobre la conveniencia e importancia de los deseos. Del mismo modo, las creencias prudenciales no son creencias morales.[39] La cuestión es que las creencias morales incluyen juicios intelectuales sobre la adecuación e importancia de los motivos de interés propio. Esto es lo que, afirmo, da

autoridad a la prudencia sobre el deseo, y da autoridad a la moral sobre la prudencia.

7.6.2 La arbitrariedad de la moral en la visión Humeana

Un humeano no puede dar una explicación similar de la autoridad de la prudencia o de la moral. Ya hemos visto, en la sección anterior, la inadecuación del relato humeano de la prudencia: el humeano debe retratar el razonamiento prudencial como si fuera, a lo sumo, un razonamiento sobre cómo satisfacer un deseo presente particular, el deseo de que los propios deseos futuros sean satisfechos. Pero no hay ninguna razón, dentro de la concepción humeana, por la que uno deba tener tal deseo, e incluso si uno lo tiene, no hay ninguna razón por la que se le deba dar un peso especial en relación con otros deseos presentes. El humeano, por tanto, no tiene en cuenta la noción intuitiva de la autoridad racional de la prudencia.

Del mismo modo, el humeano no tiene ninguna explicación de la autoridad de la moralidad, ni ninguna explicación plausible de por qué deberíamos dar a la moralidad algo parecido a la importancia que le damos. Un humeano debe representar la moralidad como derivada en última instancia de alguna especie de deseo, como el deseo por el bienestar de los demás. El propio Hume postuló la "simpatía natural" por otros seres humanos como la fuente de la moralidad.[40] Por lo tanto, toda la teorización moral y el razonamiento moral se considera que se refieren simplemente a cómo satisfacer los deseos derivados de esta tendencia psicológica particular. Pero la simpatía es sólo una fuente de emoción entre otras; no hay ninguna razón para dar un peso especial a los deseos que provoca, en contraposición a los causados por otras emociones. Por lo tanto, tendría tanto sentido dedicar un campo especial de investigación, con complejos sistemas de principios, etc., a averiguar cómo satisfacer

los deseos causados por la envidia, la lujuria o el odio, como dedicarlo a la simpatía. Si mis emociones hostiles entran en conflicto con mi simpatía natural, no hay ninguna razón humeana por la que esta última deba prevalecer generalmente. Los juicios morales no parecen corresponder más que a un subconjunto arbitrariamente seleccionado de nuestras emociones y deseos.

¿Cómo podría un humeano intentar evitar esta consecuencia? Una forma podría ser describir la moralidad como resultado, no de alguna emoción particular o subconjunto de emociones, sino de una combinación de todos nuestros deseos. Pero si esto fuera así, entonces deberíamos considerar que todas las emociones dan lugar a razones morales, y deberíamos asignar un peso moral a todos los deseos en proporción a su fuerza. Esto no es así. No consideramos que el amor y el odio proporcionen razones morales igualmente buenas. Por ejemplo, no pensamos que un odio suficientemente fuerte justifique el asesinato; en la mayoría de las circunstancias consideraríamos que esa emoción no proporciona ninguna razón para la acción, o incluso hace que las acciones impulsadas por ella sean peores.

Otro enfoque para un humeano podría ser representar la moralidad como derivada de deseos de segundo orden, como el deseo de adquirir o perder algún otro deseo. Esto podría dar sentido a la noción de que las consideraciones morales incluyen evaluaciones de otros motivos. Pero este planteamiento se enfrenta al mismo tipo de problema que el anterior. Supongamos que nos encontramos con alguien que no sólo odia a las mujeres, sino que quiere seguir odiándolas. No caracterizaríamos su deseo de segundo orden como un motivo moral, como tampoco lo es el deseo de primer orden. Además, para un humeano, el comportamiento de uno siempre está determinado por lo que quiere en el presente, no por lo que quiere querer, lo que querrá en el futuro o cualquier otra cosa. Por

tanto, relegar los juicios morales a juicios sobre lo que satisface los deseos de segundo orden impediría que influyeran directamente en las acciones. Eso sería un error; los juicios morales evalúan las acciones y los resultados en el mundo, y es así como pueden entrar directamente en conflicto con los deseos y los juicios prudenciales, y tal vez prevalecer sobre ellos.

7.6.3 El problema de la debilidad de la voluntad

La debilidad de la voluntad -fenómeno en el que un individuo sucumbe a la tentación de actuar en contra de su propio juicio- es una experiencia común. Se presenta en dos variedades. En la primera, la persona actúa por algún deseo, en contra de sus propios intereses reconocidos. Un ejemplo es el estudiante que decide pasar la noche de fiesta en lugar de estudiar, a pesar de que cree que estudiar sería lo mejor para sus intereses a largo plazo. En la segunda variedad, una persona actúa por algún deseo o creencia prudencial, en contra de sus propios valores morales. Un ejemplo sería un funcionario que acepta un soborno, a pesar de ser consciente de que hacerlo está mal.

Cuando se les habla de este fenómeno, algunas personas insisten en que es imposible que alguien no haga lo que cree mejor.[41] Supongo que esta opinión es errónea. La debilidad de la voluntad es un fenómeno comúnmente observado. La teoría se encarga de concordar con los hechos observados, y no a la inversa.

Las características más destacadas del fenómeno son las siguientes:

i) El agente considera que una acción, *A*, es la mejor.

ii) El agente se siente "impulsado" a realizar otra acción, *B*.

iii) Decimos (y, cuando el agente somos nosotros, sentimos) que la voluntad del agente es "débil" en algún sentido; que si el

agente fuera más fuerte, haría *A*. En relación con esto, es "difícil" hacer *A* pero "fácil" hacer *B*.

iv) El agente suele sentir culpa o remordimiento al hacer *B*.

v) Juzgamos que el agente es hasta cierto punto irracional al hacer *B*.

vi) Aunque el agente puede decir cosas como "no pude resistirme", el agente actúa voluntariamente. Por ejemplo, no pierde repentinamente el control de sus músculos y acaba haciendo B a modo de espasmo muscular. No intenta hacer *A* pero "falla" y hace *B* por accidente. Su psique no está controlada por una personalidad alternativa. Ninguna fuerza externa interfiere en su proceso de toma de decisiones. Aparte de la sensación de conflicto entre sus diferentes motivos, la acción del agente se produce exactamente como lo hacen normalmente las acciones intencionales. Incluso puede realizar una planificación consciente para llevar a cabo el acto *B*, como en el caso del fumador que, aunque aparentemente "intenta dejar de fumar", va deliberadamente a la tienda a comprar un paquete de cigarrillos.

He expuesto este último punto con cierta elaboración porque las víctimas de la debilidad de la voluntad tienen tendencia a poner excusas, alegando que "no pueden evitar" su comportamiento, que son "adictos" (a los cigarrillos, a los alimentos grasos, etc.), etcétera. Una comprensión filosófica del fenómeno exige resistirse a esa tendencia. No hay ninguna razón, aparte de nuestros deseos de evitar incurrir en culpa y de evitar herir los sentimientos de los demás culpándoles, para negar que las víctimas de la debilidad de la voluntad puedan controlar su comportamiento. Aparte de su paradigmática culpabilidad, las acciones de la debilidad de voluntad muestran todas las características típicas de las acciones voluntarias e intencionales.

Ahora, fíjese en las asimetrías implícitas en las características anteriores: uno tiene la tentación de hacer *B*; no tiene la tentación de hacer *A*. Hacer *A* es difícil; hacer *B* es fácil. *A* concuerda con nuestro "mejor juicio"; *B* no. Todo esto implica que las razones para hacer *A* son superiores a las razones para hacer *B*, pero el agente quiere hacer *B*. Un humeano no puede entender esto. Para el humeano, cuando dos motivos entran en conflicto, la única pregunta es, ¿qué deseo es más fuerte? Si *A* satisface un deseo más fuerte que *B*, entonces el agente hará *A*. Si, por otro lado, *B* responde al deseo más fuerte, entonces el agente debería, racionalmente, elegir hacer *B*. ¿En qué sentido sería entonces "débil" el agente? ¿En qué sentido estaría "cediendo" a una tentación? ¿Por qué sentiría culpa o remordimiento? ¿Y qué sentido tiene la idea de que una alternativa sea "más difícil" de elegir que la otra? Por supuesto, desde un punto de vista humeano, se podría decir que es "difícil" realizar la acción para la que se tienen los motivos más débiles; más exactamente, que es imposible.[42] Pero éste no es el sentido relevante de "difícil", ya que las acciones de voluntad fuerte no son imposibles..

Nuestra concepción ordinaria del fenómeno implica que, en casos de debilidad de la voluntad, los motivos en conflicto difieren en especie. Supongamos, por ejemplo, que estoy intentando decidirme entre pedir salmón o fettuccine para cenar. Supongamos que la única consideración que tengo en mente es el gusto. Y supongamos que decido que el salmón sabe mejor. Aquí no hay lugar para la "debilidad de la voluntad". Si pido salmón, no será "debilidad". Si de algún modo pido fettuccine, se trataría simplemente de una extraña disfunción psicológica, no de un ejercicio de fuerza de voluntad. Para un humeano, toda motivación es fundamentalmente una pieza: todo es cuestión de deseos presentes de diversas fuerzas. Por lo tanto, todos los casos de conflicto de motivos son similares al conflicto salmón-fettuccine.

La concepción racionalista, sin embargo, introduce tipos de motivación fundamentalmente diferentes: nuestros deseos pueden entrar en conflicto con nuestros juicios sobre qué hacer. Ésa es la esencia de la debilidad de la voluntad. Y las doctrinas de la autoridad de la prudencia y de la moralidad explican las asimetrías relativas a la "debilidad", la culpa y la irracionalidad.

Se suele considerar que el "problema de la debilidad de la voluntad" consiste en explicar cómo es posible la debilidad de la voluntad. Antes de responder a eso, preguntémonos primero: ¿por qué no sería posible? Sostengo que la concepción humeana de las razones prácticas implica que la debilidad de la voluntad es imposible, mientras que la concepción racionalista no tiene tal implicación. En un caso de debilidad de la voluntad, el agente actúa por su propia voluntad, pero sabe que tiene razones normativas suficientes para hacer otra cosa en su lugar. Hay dos características de la concepción humeana que descartan tales casos. En primer lugar, como vimos en la sección 7.5, la concepción humeana implica que los factores que determinan la fuerza normativa de las razones son los mismos que los que determinan la fuerza motivadora de las razones. Por tanto, sería imposible que lo que un agente está motivado a hacer se desprenda de lo que es racional que haga, como ocurre en los casos de debilidad de la voluntad..

En segundo lugar, según el punto de vista humeano, todas las motivaciones pueden compararse a lo largo de una única dimensión, en la que esa dimensión es lo único relevante para su capacidad de inducir a la acción. Desde el punto de vista humeano, es la fuerza del deseo relevante lo que determina su capacidad para inducir a la acción, y todos los deseos pueden compararse en términos de su fuerza.[43] La explicación humeana de la motivación, por tanto, sugiere fuertemente una doctrina del determinismo, es decir, que en cualquier situación dada (teniendo en

cuenta las motivaciones del agente), sólo hay una cosa que un agente dado puede hacer. Y esto parece, al menos a primera vista, entrar en conflicto con la observación de que las acciones de voluntad débil se eligen libremente.[44] Más concretamente, parece que cuando se describe un escenario en el que una persona experimenta motivaciones contradictorias y se ve tentada a actuar en contra de su buen juicio, cualquiera de las dos decisiones es posible; sea cual sea el curso de acción que elija el agente, parece que podría haber habido una situación en la que una persona tuviera las mismas razones disponibles pero eligiera el otro curso de acción.

Un humeano podría tratar de explicar esto proponiendo que en tales casos, el agente tiene motivaciones exactamente igual de fuertes para cada curso de acción. Pero si la fuerza del deseo varía continuamente, entonces la probabilidad de que dos deseos distintos tengan exactamente la misma fuerza es cero;[45] por tanto, no parece razonable creer que ésta sea la explicación de la debilidad de la voluntad. Y, de nuevo, la teoría del "empate exacto" no explicaría las diversas asimetrías implícitas en la noción de debilidad de la voluntad, como el hecho de que se diga que un curso de acción y no otro está de acuerdo con el mejor juicio de uno, que es racional, que es difícil, etc.

La concepción racionalista de las razones no se enfrenta a los problemas a los que se enfrenta la concepción humeana, ya que la concepción racionalista rechaza los dos principios problemáticos de la visión humeana. En primer lugar, rechaza la idea de que la fuerza normativa de las razones esté determinada por los mismos factores que determinan la fuerza motivadora. La fuerza normativa de una razón para la acción viene determinada, más bien, por los valores objetivos -por ejemplo, cuánto bien produciría una acción dada (o cuánto se cree que produciría). En esta teoría no se sugiere que deba haber alguna

dimensión psicológica de fuerza motivadora a la que esto corresponda siempre. En segundo lugar, rechaza la idea de que todas las motivaciones puedan compararse a lo largo de una única dimensión de "fuerza". Más bien, mientras que los deseos pueden compararse en términos de fuerza, nuestros juicios normativos son un tipo de motivación cualitativamente diferente. Los juicios no tienen una dimensión psicológica que pueda compararse con la fuerza de un deseo. Existe, por supuesto, la "fuerza" de una creencia (es decir, el nivel de certeza del creyente), y existe el grado de valor objetivo que uno cree que tiene algún objetivo, pero ninguna de estas cosas es conmensurable con la fuerza de los deseos. En consecuencia, la concepción racionalista no puede sustentar una imagen determinista de la toma de decisiones humana. La inconmensurabilidad psicológica de los deseos y las creencias normativas puede explicar por qué, en casos de debilidad de la voluntad, pensamos que cualquiera de las dos decisiones es posible.

Ahora quiero aclarar dos cosas sobre esta explicación. En primer lugar, no afirmo que la concepción racionalista de las razones prácticas implique la existencia de debilidad de la voluntad. Es teóricamente posible combinar la concepción racionalista de las razones con una visión en la que, en casos de conflicto entre diferentes tipos de motivos, la gente siempre hace lo que percibe que es moralmente necesario -o una en la que siempre hace lo que espera que satisfaga mejor sus deseos. La concepción racionalista no implica ni el indeterminismo ni la existencia del libre albedrío, ni explica por sí misma "cómo funciona el libre albedrío". No tengo nada que ofrecer sobre cómo funciona el libre albedrío. Lo que sostengo es que la concepción racionalista es superior a la concepción humeana en el sentido de que la concepción racionalista no implica que la debilidad de la voluntad sea imposible, como lo hace la concepción humeana.

En segundo lugar, es importante entender por qué, al permitir los casos de debilidad de la voluntad de la forma que he sugerido, no estoy dando marcha atrás en la tesis de la sección 7.2. Sigo manteniendo que el reconocimiento por parte de un agente de razones normativas produce necesariamente al menos cierta motivación para actuar de acuerdo con esas razones. Sigo manteniendo que el reconocimiento de razones normativas por parte de un agente produce necesariamente al menos cierta motivación para actuar de acuerdo con esas razones. Recordemos que en la sección 7.2 había dos argumentos centrales. El primero se basaba en la idea de que siempre debe ser posible que un agente haga lo que racionalmente debería hacer. El segundo se basaba en la idea de que está justificado a priori interpretar que otras personas, si es posible, son en general racionales. Estos argumentos no exigen una explicación determinista de la elección humana, ni implican que la fuerza de una razón normativa deba corresponderse con la fuerza del motivo del agente en ningún sentido.

7.6.4 Cómo el antirrealismo socava la moralidad

Ahora estamos en condiciones de ver cómo la creencia en el antirrealismo socava la moralidad. Socava tanto nuestras creencias morales como nuestra motivación para comportarnos moralmente. Este hecho no demuestra que el antirrealismo sea falso. Pero muestra que, si el antirrealismo es falso, también es pernicioso.

Algunos de mis colegas antirrealistas parecen encontrar esta tesis sorprendente y difícil de aceptar. Antes de abordar lo que dicen, puede poner las cosas en perspectiva observar que no sólo yo y otros realistas creemos esto; incluso algunos antirrealistas prominentes han admitido de hecho que su trabajo socava la moralidad, aunque no lo expresan del todo así. Justo antes de atacar la noción de valor objetivo, J. L. Mackie

admite que "los juicios morales ordinarios incluyen una pretensión de objetividad, una suposición de que existen valores objetivos justo en el sentido en el que yo me ocupo de negarlo".[46] De ello se deduce que su propio libro sirve para socavar las creencias morales ordinarias. Mackie puede creer que puede poner algo mejor en su lugar,[47] pero esto es bastante dudoso por las razones discutidas en esta sección. Pero es cuando los antirrealistas llegan a explicar por qué otras personas creen en la moralidad objetiva cuando revelan más claramente cómo sus propias teorías, de ser aceptadas, socavarían la motivación moral. Consideremos estas explicaciones proporcionadas, respectivamente, por los filósofos antirrealistas J. L. Mackie y Michael Ruse:

> *Necesitamos la moral para regular las relaciones interpersonales, para controlar algunas de las formas en que las personas se comportan entre sí, a menudo en oposición a inclinaciones contrarias. Por tanto, queremos que nuestros juicios morales tengan autoridad tanto para otros agentes como para nosotros mismos: la validez objetiva les daría la autoridad necesaria.*[48]

> *La cuestión de la moral... es que es una adaptación para conseguir que vayamos más allá de los deseos, anhelos y temores habituales, y que interactuemos socialmente con la gente. ¿Cómo consigue que hagamos esto? Llenándonos de pensamientos sobre obligaciones y deberes, y así .... [E]ntonces nos mueve a la acción, precisamente porque pensamos que la moral es algo que se nos impone.... Si la moral no tuviera este aire de exterioridad u objetividad, no sería moral y (desde una perspectiva biológica) no haría lo que pretende.*[49]

De ello se desprende que los antirrealistas, al atacar su objetividad, trabajan para socavar estas importantes funciones de la moral.

La cuestión también está clara, supongo, por lo que he dicho

anteriormente. La única razón que los antirrealistas pueden ofrecer para explicar por qué debemos comportarnos moralmente es que hacerlo ayudará supuestamente a satisfacer algún deseo preexistente nuestro, como el deseo de ayudar a otras personas. Si descubrimos que no tenemos este deseo, o si descubrimos que algún otro deseo es más fuerte en ese momento, entonces, según el relato Humeano, sería irracional ser moral. En cualquier caso, la apelación a las exigencias morales no tendría ninguna fuerza especial, en relación con cualquier otra apelación al deseo. Además, si tengo razón en algunas de mis afirmaciones centrales en este libro, tanto sobre los significados de las exigencias morales como sobre cómo nos motivan a actuar, entonces el antirrealismo socava la motivación moral de una manera muy directa: ataca las creencias que constituyen nuestras razones para actuar moralmente.

Muchos antirrealistas se opondrían a esta conclusión. Dirían que la motivación moral proviene enteramente de las emociones, por lo que lo que se necesita para reforzar la moralidad es simplemente entrenar a las personas (aplicando influencias no cognitivas) para que tengan sentimientos sociales de simpatía más fuertes; y un antirrealista puede hacerlo tan bien como un realista. Yo digo que las emociones no son suficientes. Las personas a menudo no tienen las emociones que el moralista humeano quiere que tengan, y puede que no tengan ninguna razón humeana (es decir, ninguna razón proporcionada por sus deseos actuales) para hacer algo para cambiar esa situación -pueden, por ejemplo, no tener ninguna razón humeana para someterse a la instrucción moral o a la exhortación.

Consideremos dos ejemplos. En primer lugar, supongamos que acabamos de leer el argumento de Peter Singer para demostrar que tenemos la obligación moral de ayudar a los pobres.[50] Así, según Singer,

cada uno de nosotros está obligado a donar gran parte de sus ingresos a organizaciones benéficas que luchan contra la pobreza en el mundo. Supongamos que el argumento de Singer me parece convincente, considerado desde el punto de vista moral. Pero entonces reflexiono que, si realmente hiciera lo que es "moralmente necesario" según el argumento, me privaría de muchas cosas que actualmente disfruto, como salir a comer, ir al cine y otros lujos. Me gustan mucho esas cosas. Así que prefiero no tener que regalar mi dinero. Así que se me ocurre preguntarme: claro, donar a la caridad es "moralmente obligatorio", pero ¿realmente tengo que ser moral? Me dirijo al campo de la Metaética, donde encuentro que la mayoría de los expertos proclaman que no hay requisitos morales objetivos. Reflexiono que este "requisito moral" que propone Singer es sólo una cuestión de convenciones sociales (como dicen los relativistas culturales), o es sólo algo que me produce un tipo de emoción particular (como dicen los no cognitivistas y los subjetivistas individualistas), o es sólo una ilusión (como dicen los nihilistas). Creyendo una de estas cosas, ¿por qué iba a obedecer el requisito moral? ¿Por qué iba a sacar mi chequera y empezar a donar ese dinero, sabiendo que mis propios intereses se verán gravemente perjudicados al hacerlo? La única razón que Singer puede ofrecerme es que, actuando moralmente, ayudaré a satisfacer un deseo mío, y quizás me haga más feliz.[51] Esto es patéticamente inadecuado a la luz del tipo de sacrificio que Singer está pidiendo. Por supuesto, tengo algunos sentimientos hacia los pobres del Tercer Mundo: lamento que se estén muriendo de hambre y siento cierta satisfacción cuando envío dinero para ayudarlos, pero ni por asomo estos sentimientos son lo suficientemente fuertes como para llevarme a regalar, digamos, 30.000 dólares al año. En realidad, el placer que me produce el envío de dinero a Oxfam no es obviamente mayor -de hecho, parece mucho menor- que el que me produce, por ejemplo, la lectura de una buena novela.[52] Y la novela sólo cuesta unos 8 dólares.

Entonces, ¿por qué iba a regalar miles de dólares a la caridad? Es cierto que mis deseos no se limitan a disfrutar, pero la cuestión sigue siendo que mi deseo de ayudar a los pobres no es tan fuerte como muchos de los deseos con los que compite. Una vez que me "dé cuenta" de que la única razón para obedecer los requisitos morales sería ayudarme a satisfacer mis preferencias personales, seguramente me sentiré con derecho a hacer el juicio de que, en general, preferiría no ser moral, y a actuar en consecuencia. Que el filósofo de la moral me diga que simplemente no me doy cuenta de la fuerza de mis sentimientos de simpatía por otras personas, etc., me parecerá tan creíble como si alguien que no me conoce me dijera que yo preferiría ser bombero que profesor.

En segundo lugar, supongamos que Jon, un hombre casado, está considerando la posibilidad de engañar a su mujer. Pide consejo a su mejor amigo, Sam, y se produce el siguiente diálogo:

Jon: [Después de explicar la situación:] Entonces, ¿crees que debería engañar a Kate?

Sam: Kate se quedaría destrozada si supiera que estás considerando esto.

Jon: No te preocupes, no se va a enterar. [Jon discute cómo va a asegurarse de que Kate no lo sepa].

Sam: ¿No crees que está mal engañar a tu mujer?

Jon: Oh, eso. Eso es sólo una convención arbitraria que la gente inventó. Sé que a mucha gente no le gusta la infidelidad conyugal, pero ¿por qué debería hacerles caso?

Sam: ¿No te sientes un poco culpable por pensar en traicionar a Kate, después de todo lo que habéis pasado juntos?

Jon: Bueno, me sentí culpable durante un tiempo. Pero luego

> me di cuenta de que era estúpido sentirse culpable. Quiero decir, ¿quién dice que la gente tiene que ser fiel a sus esposas? Eso es sólo una regla arbitraria que alguien inventó. Si no quiero ser fiel, ¿por qué debería serlo? Admito que, de todas formas, puede que experimente una culpa irracional, pero creo que aún así valdrá la pena por toda la diversión que voy a tener.

Jon ha aprendido bien las lecciones de sus clases de filosofía. Puede que no haya seguido todas las sutilezas y complicaciones de las distintas versiones del antirrealismo, pero ha captado lo esencial: nos inventamos requisitos morales, impulsados por algún subconjunto arbitrariamente seleccionado de nuestras emociones y deseos. Nuestro sentido moral (no tiene autoridad sobre los que disienten de sus recomendaciones o incluso sobre nosotros si nos inclinamos a cambiar de opinión".[53]

No creo que mis dos ejemplos sean inusuales. Tampoco veo ningún fallo fundamental ni en el razonamiento de Jon ni en el mío en los ejemplos anteriores, aparte del error de no reconocer los valores objetivos.

A menudo oímos quejas sobre la decadencia de la moral en la sociedad moderna. No sé si estas quejas son justas; no tengo estadísticas sobre la frecuencia de los actos inmorales en distintas épocas. Pero la aceptación generalizada del antirrealismo en nuestra cultura no puede ser buena para la moral. Seguramente habrá personas que se desentiendan de muchas de las exigencias de la moral, una vez que decidan que la moral es irreal. Por supuesto, el intuicionismo ético no puede garantizar un comportamiento moral, como tampoco puede garantizar unas creencias morales correctas. Ninguna teoría filosófica puede evitar que la gente se entregue a la debilidad de la voluntad. Pero, al menos, la teoría no fomenta positivamente tales lapsos al revestirlos con el manto de la racionalidad práctica.

7.7 ¿Por qué ser moral?

¿Por qué debemos comportarnos moralmente? La respuesta del intuicionista ético es la siguiente. Nuestras intuiciones nos justifican *prima facie* para adoptar ciertas creencias morales. Siempre que estas creencias morales sobrevivan a cualquier razón para dudar en ellas, están justificadas *tout court*. Estas creencias constituyen en sí mismas razones para la acción, en virtud de su contenido evaluativo. No hay más preguntas que responder acerca de por qué uno debería estar motivado por evaluaciones correctas; es simplemente la naturaleza de la evaluación la que responde a preguntas sobre qué hacer. Aceptar un juicio evaluativo es juzgar algo como una razón para la acción: aceptar el juicio de que robar está mal, por ejemplo, es aceptar la propiedad de *x* por parte de otra persona como una razón normativa para no tomar *x* sin su permiso.[54]

Algunos filósofos, con los que tengo mucha simpatía, sostienen que "¿Por qué debo ser moral?" es una pregunta sin sentido.[55] Consideremos tres interpretaciones de la pregunta:

i) "¿Por qué debo hacer lo que moralmente debo hacer?" Esto parece significar: "¿Por qué debo hacer lo que debo hacer?", que es una pregunta sin sentido.

ii) "¿Por qué debería hacer las cosas que se consideran generalmente, y que quizás yo mismo he considerado hasta ahora, morales?". En este caso, la respuesta parece ser: "Porque esas creencias morales son, en general, verdaderas".[56] Si se pregunta por qué hay que aceptar esas creencias, la respuesta es que está justificado creer en el contenido de nuestras intuiciones éticas, en ausencia de razones para dudar de ellas, tal como se discute en el

capítulo 5.

iii) "¿Por qué es racional ser moral?" Esta interpretación ofrece la mejor esperanza de dar sentido a la pregunta. Si esto es lo que quiere decir, la respuesta es que los juicios morales son en sí mismos productos de la razón; son evaluaciones intelectuales de acciones, estados de cosas, etcétera. Además, son evaluaciones del tipo que tienen en cuenta todas las razones relevantes, incluidas las razones prudenciales y los deseos actuales, como se expone en la sección 7.6.1. Por lo tanto, lo que juzgamos moral es lo que nos parece correcto. Así pues, lo que juzgamos moral es lo que la razón evalúa como digno de elección.

Algunos lectores pueden quedar insatisfechos con mis respuestas, por una de estas dos razones. Por un lado, algunos pueden pensar que, a pesar de mi respuesta a la pregunta (*iii*), sigue existiendo la pregunta de por qué hay que ser racional. Otros pensarán que, a pesar de lo que he dicho al respecto, la pregunta (*i*) parece una pregunta real. Por supuesto, suena mejor si se sustituye "lo que debería hacer" por "lo que es moralmente correcto" y si se elimina "debería hacer", así: "¿Por qué hacer lo que es moralmente correcto? Estoy de acuerdo en que, a primera vista, parece una pregunta con sentido. Ahora quiero dar una razón para pensar que esta apariencia es ilusoria, y que ni "¿Por qué ser racional?" ni "¿Por qué hacer lo que es moralmente correcto?" es una pregunta legítima.

Para que estas preguntas tengan "sentido" y sean "serias" en el sentido previsto, debe haber, en primer lugar, algo que, en el contexto actual, cuente como una respuesta satisfactoria a una pregunta del tipo "¿Por qué hacer *x*?". Esta respuesta diría presumiblemente algo que es cierto para *x* pero que no es cierto para otras acciones que no deberían realizarse. Supongamos que esa "cosa" es *F*. En otras palabras, supongamos que "Porque *x* es *F*" es una respuesta satisfactoria a la

pregunta filosófica general "¿Por qué hacer *x*?". No estoy asumiendo que "¿Por qué ser moral?" o "¿Por qué ser racional?" tenga una respuesta satisfactoria; no asumo que las acciones morales o las acciones racionales tengan *F*. Sólo asumo que hay algo que podría decirse sobre estas acciones que, de ser cierto, contaría como respuesta a esas preguntas. Esto parece un requisito razonable para imponer a las preguntas que debemos tomar en serio.

Sin embargo, incluso si existiera tal propiedad *F*, algunos filósofos podrían seguir diciendo: "Sí, pero ¿por qué hacer las cosas *F*?". Su pregunta estaría ciertamente bien formada gramaticalmente; es más, probablemente sonaría, al menos para muchos, como si fuera una pregunta con sentido. Pero, por la naturaleza del caso, no podría ser una pregunta legítima; por hipótesis, *F* es la propiedad que responde a las preguntas "por qué hacer tal cosa". Por tanto, una vez que se ha dicho que una acción es *F*, no puede haber ninguna otra pregunta sobre por qué hay que hacerla. A la inversa, si hay que responder a alguna pregunta más, entonces "Porque *x* es *F*" no habría sido una respuesta suficiente a la pregunta de por qué hay que hacer *x*.

Por lo tanto, podemos predecir que habrá algunas preguntas del tipo "¿Por qué hacer las cosas que tienen la característica *F*?" que superficialmente parecen legítimas pero no lo son. Se trataría de preguntas en las que el rasgo citado ya constituye intrínsecamente (o conlleva la existencia de) una razón para la acción. ¿Cuáles serían los candidatos más plausibles para estas preguntas? Parece que los candidatos más obvios para *F* son "moral" y "racional" (más de un rasgo puede desempeñar el papel de *F*); si algo contara como inherente a la existencia de una razón adecuada para hacer *x*, sería que *x* es correcto, o que *x* es racional.[57]

El único otro candidato obvio para *F* es la propiedad de satisfacer los

propios deseos. Pero, ¿no podría preguntar: "¿Por qué debería satisfacer mis deseos?"? A esto, un humeano respondería que la pregunta es ilegítima, porque mi deseo de algo ya constituye una razón para que lo persiga. Pero, ¿no está igual de claro que el hecho de que algo sea bueno ya constituye una razón para que lo persiga? En todo caso, esto parece más claro. Supongamos que tengo deseos que yo mismo considero malos. Por ejemplo, siento un impulso repentino e inexplicable de saltar de un edificio. ¿Es realmente obvio que entonces tengo una razón para saltar del edificio, es decir, algo que haría que saltar del edificio fuera a primera vista racional?[58] ¿Es obvio, por otra parte, que tengo una razón para no saltar del edificio? Por supuesto, puedo tener tanto una razón a favor como en contra. La cuestión es que es más obvio que tengo una razón para no saltar por saber que saltar del edificio sería malo, que que tengo una razón para saltar por sentir el deseo de saltar.

Llego a la conclusión de que preguntas como "¿Por qué ser moral?" y "¿Por qué ser racional?" se encuentran entre los candidatos más claros a las preguntas ilegítimas de "por qué". Son como la pseudopregunta "¿Por qué pesan 10 libras más que 5 libras?".

Notas

1. Utilizo el término "razones normativas" porque considero que el concepto de racionalidad es un concepto normativo central. Considero que las razones morales son un subconjunto propio de las razones normativas.

2. Hume ([1739] 1992, pp. 413-17). Más precisamente, en (a), debemos leer "que crees que *A* logrará o podría lograr" como "tal que tu probabilidad subjetiva de que ocurra dado que haces *A* es mayor que tu probabilidad subjetiva incondicional de que ocurra". En (b), deberíamos leer "que tienes razones para creer que *A* se cumplirá o podría cumplirse" como "tal que la probabilidad epistémica, en tu evidencia total, de que ocurra dado que haces *A* es mayor que la probabilidad epistémica incondicional, en tu evidencia total, de que ocurra".

3. El argumento se deriva de Hume ([1739]1992, pp. 413-18, 455-70) y es discutido ampliamente por Michael Smith (1994, esp. p. 12). Es uno de los argumentos más populares de toda la filosofía y ha sido respaldado por Mackie (1977, pp. 40-1), Gibbard (1990, p. 33), Hare (1999, p. 1), NowellSmith (1954, pp. 39-42), Joyce (2001) y Fumerton (1990, capítulo 3), entre otros.

4. Kant ([1785] 1959) y Nagel (1970) rechazan la concepción humeana. McDowell (1978) rechaza la brecha creencia/deseo. Ross ([1930] 1988, pp. 157-8, 163) rechaza y Frankena (1958) pone en duda el Magnetismo de los Valores.

5. Considero que esto es válido para cualquier sentido del "debería", no sólo para el "debería" moral.

6. Tener una motivación disponible para hacer *A* es tener un estado mental (posiblemente disposicional o subconsciente) tal que, o bien le motiva a uno a hacer *A*, o bien puede hacer que le motive.

7. Una variante es la idea de que debemos in terpretar a los demás como si tuvieran creencias generalmente verdaderas (Davidson 1990, pp. 129-30); sin embargo, esta versión del principio de caridad es inferior a la que yo discuto. Está justificado que atribuyamos a los demás creencias falsas, siempre y cuando sea lo más lógico para ellos tener esas creencias falsas. También está justificado atribuir a otros declaraciones falsas (incluidas las mentiras), cuando lo más racional (quizás por razones prácticas) sea que hagan esas declaraciones falsas.

8. El término "placer físico" es un término erróneo: los placeres sensoriales no son más físicos que los placeres emocionales. Si el fisicalismo en la filosofía de la mente es cierto, ambos son físicos; si el dualismo es cierto, ambos son no físicos.

9. Bentham ([1781] 1996, p. 11); Hume ([1739] 1992, p. 414). Compárese con la afirmación de von Mises (1996, pp. 13-15) de que toda acción tiene como objetivo eliminar el malestar sentido y lograr la felicidad.

10. Compárese la observación, en la sección 5.6, de que nuestras percepciones e intuiciones se centran en el mundo. El error de considerar que la conciencia se dirige principalmente a los estados internos es similar al error de considerar que nuestros objetivos se refieren principalmente a los estados internos.

11. Hubin (1980, p. 75) despliega un ejemplo similar, para hacer el punto relacionado de que el interés propio no consiste en la satisfacción del deseo. Véase también el ejemplo de la "máquina de la experiencia" de Nozick (1974, pp. 42-5), que muestra que

los objetos de nuestros deseos no suelen ser estados internos. La mayoría de las personas no se enchufarían a la máquina por la misma razón por la que no venderían a sus hijos como esclavos: eso no les daría lo que quieren.

12. Compárese la clásica refutación del egoísmo psicológico de Butler ([1729] 1964, pp. 20--5).

13. Michael Smith (1994, p. 114) trata los deseos de esta manera. Nagel (1970, p. 29) sugiere que podemos atribuir un deseo a alguien en virtud del hecho de que está motivado de una manera determinada.

14. Esto no tiene por qué equivaler a lo que Michael Smith (1994, pp. 104-111) denomina "concepción fenomenológica del deseo", ya que no es necesario suponer que los apetitos y las emociones deban tener siempre qualia ocurrentes. Compárese: podemos identificar introspectivamente nuestras creencias, aunque según la visión recibida, las creencias no tienen qualia.

15. Pocos de los filósofos de esta corriente utilizarían la palabra "intuición", o incluso la palabra "obvio", al hacer este llamamiento. Más bien, se limitarían a afirmar la concepción humeana sin argumentos, y es esto lo que me lleva a interpretar que se basan en la intuición.

16. Hume ([1739]1992, p. 458).

17. El ejemplo es de Anscombe (1963, pp. 56-7). La terminología de "dirección de ajuste" es de Searle (1979, pp. 3-4; 1983, pp. 7-9), que puede haberla obtenido de J. L. Austin.

18. Michael Smith (1994, pp. 111-16).

19. Parafraseado de Michael Smith (1994, p. 116). En realidad, Smith dice que cualquier estado mental con dirección de ajuste mundo-mente es un deseo y, por tanto, que la motivación requiere un deseo. He debilitado ligeramente las afirmaciones de Smith para dar al argumento la mejor oportunidad de éxito.

20. Resulta interesante que Smith diga que (2) es "inatacable" y que sólo (3) puede ser cuestionada.

21. Una lectura normativa de (2) es sugerida por el uso de Smith de la frase "estado con el que el mundo debe encajar" (1994, p. 116; el énfasis es mío). Sin embargo, dado que el argumento se refiere a razones motivadoras, más que a razones normativas, parece que se requiere una lectura descriptiva. Esta lectura se ve apoyada por la observación de

Smith de que un estado con dirección de ajuste mente-mundo "tiende a desaparecer" cuando uno percibe que el mundo no se ajusta a él (p. 118). Shafer-Landau (2003, p. 35) argumenta que la noción de dirección de ajuste debe entenderse normativamente, del siguiente modo: un estado con dirección de ajuste mente-mundo es aquel que tiende a desaparecer ante la percepción de que su contenido no está satisfecho. Nótese que el equívoco del que acuso a Smith en el texto surgiría tanto si se toma una interpretación normativa como descriptiva de la dirección de ajuste.

22. Scanlon (1998, pp. 33-5) hace una afirmación similar, en el sentido de que un juicio normativo puede constituir una motivación, lo que considero equivalente a la afirmación de que hacer un juicio normativo puede constituir tener una meta. La noción de "tener una meta" podría, en cambio, considerarse que requiere una intención1 de perseguir la meta, pero en esa interpretación, el argumento de Smith fallaría, ya que la presencia de una razón motivadora para perseguir x no implica la presencia de una intención de perseguir x.

23. Hume ([1739J 1992, p. 458).

24. Burne ([1739]1992, p. 416). Como confirma el contexto, el punto de Hume es que ninguna preferencia puede ser criticada racionalmente.

25. "Racional" es en sí mismo un término evaluativo, sobre el que un humeano podría adoptar una teoría antirrealista (Gibbard 1990, pp. 6-7). Es natural preguntarse cómo afectaría esto a mis argumentos posteriores. La respuesta es que no debería afectarlos, siempre que la opinión antirrealista en cuestión sea plausible. Porque el antirrealismo sobre los valores es plausible sólo si permite la posibilidad de un razonamiento convincente, del tipo que implica nuestro discurso actual, sobre cuestiones evaluativas de primer orden. Mis argumentos en las siguientes secciones sobre lo que es prácticamente racional, entonces, serían simplemente instancias de tal razonamiento evaluativo.

26. Brandt (1996, pp. 51-3).

27. Brandt (1979, pp. 14-15); Williams (1981, pp. 104-5).

28. Michael Smith (1994, pp. 159-60). Smith se considera a sí mismo un antihumeano sobre las razones normativas; yo, sin embargo, lo llamo neohumeano.

29. Williams (1981, pp. 106-9).

30. Nagel (1970, pp. 39-41) considera que este punto revela una grave deficiencia en el relato humeano de la racionalidad práctica.

31. Si un gran conjunto de creencias es inconsistente, cada creencia individual puede ser lo suficientemente probable como para que uno no deba renunciar a ella. Véase Foley (1979).

32. Supongamos, por ejemplo, que el ateo logra mostrar que el concepto de omnipotencia es incoherente, que la omnipotencia es inconsistente con la omnibenevolencia, o algo parecido.

33. Stuart Rachels (comunicación personal) sugiere que podría ser razonable sólo desear que el hambre en el mundo termine, en lugar de desear que termine. Tal vez exista un sentido de "deseo" en el que tener un deseo implica algo así como la creencia de que el objeto del deseo tiene una posibilidad razonable de ser alcanzado. Pero aquí no utilizo "deseo" en ese sentido. Utilizo "deseo" en un sentido amplio que incluye los deseos. De forma similar, mi sentido de "deseo" también permite desear lo que ya se tiene.

34. Bonjour (1985, pp. 169-79). Véase Olsson (2005) para las críticas a la coherencia por su incapacidad de asegurar la probabilidad de verdad.

35. Considero que la justificación epistémica es una cuestión de probabilidad de verdad, en un sentido de "probabilidad".

36. Ryan (2003) defiende esta posición de forma persuasiva, aunque aquí está nadando contra la corriente.

37. Scanlon (1998, pp. 62-3). Me he tomado algunas libertades: donde yo hablo de "juicios evaluativos", Scanlon habla de "conclusiones sobre las razones" (que él considera el tipo más fundamental de juicio evaluativo). Esta distinción no afecta a las críticas que le hago en el texto.

38. Asumo que "*x* es lo que hace verdadero a *y*" es transitivo, de modo que si *B* es lo que hace verdadero a *A*, y *C* es lo que hace verdadero a *B*, entonces *C* es lo que hace verdadero a *A*. Las condiciones se vuelven cada vez más complejas a medida que avanza la serie, ya que cada condición menciona a la anterior.

39. A menos que el egoísmo ético sea verdadero.

40. Hume ([1751]1975, pp. 285-6).

41. Estos individuos tienen al menos a Platón de su lado. Véase el Mello (77d-e) y el Protágoras (345d).

42. Hume ([1739] 1992, pp. 399-407) rechaza el libre albedrío, como, creo, debería

hacerlo cualquiera que sostenga su visión de la motivación.

43. También se podría introducir, como segundo factor que determina la fuerza de las motivaciones, la confianza de la creencia del agente en que una determinada acción satisfará un determinado deseo. Sin embargo, el argumento expuesto en el texto seguirá siendo válido, siempre que los dos factores puedan combinarse en uno solo, por ejemplo, multiplicando el grado de creencia por la fuerza del deseo, como en los cálculos estándar de utilidad esperada.

Richard Fumerton (comunicación personal) sugiere que otras características de los deseos, además de su fuerza, podrían ser relevantes para su eficacia motivacional. Tal vez los deseos de placeres inmediatos tiendan a tener más fuerza motivacional que los deseos de beneficios lejanos debido simplemente a su contenido, más que a su fuerza. Aunque esta es una posibilidad teórica, parece falsa desde el punto de vista introspectivo: los deseos de placeres inmediatos se sienten más fuertes que los deseos de beneficios comparables pero más lejanos, y esta parece la explicación más natural de por qué los primeros deseos suelen tener más fuerza motivacional.

44. Hay una amplia literatura sobre si ese conflicto es genuino, en la que no puedo entrar aquí. Véase mi (2000b) para más información.

45. Si una variable tiene una infinidad continua de valores posibles, entonces su probabilidad de tomar cualquiera de esos valores debe ser cero; de lo contrario, la probabilidad total de todas las posibilidades sería infinita. Por tanto, si hay un continuo de muchos grados posibles de fuerza, entonces la probabilidad de que cualquier deseo tenga un grado exacto dado de fuerza debe ser cero. En consecuencia, la probabilidad de que dos deseos diferentes tengan el mismo grado de fuerza (suponiendo que no haya una conexión lógica entre ellos que les obligue a tener el mismo grado de fuerza) es cero.

En lugar de decir que las probabilidades son cero, algunas personas desearían plantear aquí probabilidades "infinitesimales", donde un número infinitesimal es uno que es menor que cualquier número real pero mayor que cero. Yo no creo en tales cantidades. Sin embargo, considero que el hecho de que una teoría implique que algo que ocurre con bastante frecuencia tiene una probabilidad infinitesimal de ocurrir sería una evidencia extremadamente fuerte en contra de esa teoría, aunque quizás no tan fuerte como lo sería si la teoría implicara que el evento tiene probabilidad cero.

46. Mackie (1977, p. 35).

47. Mackie (1977, capítulo 5).

48. Mackie (1977, p. 43).

49. Ruse (1998, p. 253). Compárese con Joyce (2001, capítulo 6). Ruse continúa expresando fuertes dudas de que podamos escapar de la influencia de la moral. No sé cómo reconciliar esto con sus observaciones sobre la importancia del sentido de la objetividad para la función de la moralidad, a menos que dude (a pesar de sus propios argumentos) de que podamos aceptar el antirrealismo.

50. Véase Singer (1993, capítulo 8); Unger (1996).

51. Singer (1993, capítulo 12) de hecho argumenta de esta manera, varios capítulos después de presentar su argumento moral.

52. Hay que reconocer que nunca he regalado las enormes cantidades de dinero que Singer exige. Pero anticipo que el disfrute derivado de ello no sería mucho mayor que el derivado de donar cantidades más pequeñas.

53. Mackie (1977, p. 106) dice esto sobre nuestro "sentido de la justicia".

54. Dancy (1993, p. 29) y Scanlon (1998, pp. 33-5) defienden este punto.

55. Véase Butchvarov (1989, pp. 48-52). La opinión se atribuye tradicionalmente a Prichard, aunque la atribución es difícil de justificar; véase su (1957, pp. 111-12).

56. No pretendo aquí descartar que la filosofía moral pueda descubrir que las creencias morales tradicionales requieren una revisión significativa. De hecho, creo que la filosofía moral ha descubierto precisamente eso. Pero, en aras del argumento, supongamos que la pregunta "¿Por qué ser moral?" se formula con respecto a un sistema moral que ya ha sido sometido al tipo de reflexión y crítica que un intuicionista sofisticado aprobaría.

57. Fumerton (1990, p. 85) hace una observación relacionada: "¿Por qué debería hacer *x*?" es una petición de una razón para hacer *x*. Pero "*x* es racional" sólo significa que *x* es lo que uno tiene más razones para hacer. Así que "¿Por qué debería ser racional?" equivale a una extraña petición de una razón para hacer lo que uno tiene más razones para hacer.

58. Quinn (1993, págs. 236-7, 246-7) hace una afirmación similar, argumentando que un mero impulso ajeno a cualquier tendencia a ver algo como bueno no proporciona una razón para la acción. Scanlon (1998, p. 38) niega que un impulso de este tipo pueda considerarse un deseo genuino.

8

Otras objeciones

En los últimos tres capítulos, he abordado tres de las objeciones más prominentes al intuicionismo, las relativas a la capacidad del intuicionismo para dar cuenta del conocimiento moral, su capacidad para explicar y resolver el desacuerdo moral, y su capacidad para dar cuenta de la motivación moral. En este capítulo, me ocuparé de otras objeciones diversas.

8.1 El argumento de la rareza

Empiezo con la famosa afirmación de J. L. Mackie de que las propiedades morales objetivas serían "raras".[1] Esta objeción se expresa en la conversación más a menudo de lo que se pone por escrito. Los filósofos a menudo descartan el realismo moral, el intuicionismo y muchos otros puntos de vista como demasiado "raro", "espeluznante" o "misterioso" para ellos.[2] Esta puede ser la principal fuente de resistencia al intuicionismo. No importa la fuerza con la que exponga el resto de los argumentos en este libro, el intuicionismo nunca será aceptado mientras los filósofos lo consideren "raro".

Mackie tenía algunas cosas específicas que decir sobre cómo las propiedades morales serían raras que no discutiré en esta sección, ya que se tratan en otra parte.[3] Aquí, mi atención se centra en su afirmación general de que las propiedades morales serían propiedades "de un tipo muy extraño, totalmente diferentes de cualquier otra cosa en el universo".[4]

Los antirrealistas pueden tener razón al calificar los valores objetivos de extraños. Pero, ¿cómo podría esto ser una razón para negar su existencia? Tal vez el argumento sea:

1. Los valores objetivos serían extraños.

2. Si algo es raro, entonces (probablemente) no existe.

3. Por lo tanto, los valores objetivos (probablemente) no existen.

¿Qué significa que algo sea "raro"? Esto podría parecer una pregunta frívola, si no fuera por la enorme influencia que el sentido de "rareza" tiene en las posiciones que se respetan e investigan seriamente en la filosofía. Consideremos cuatro interpretaciones de la noción:

Primera: Quizás decir que un tipo de cosa es rara es decir que es contraintuitivo que existan cosas de ese tipo. Pero Mackie no acepta la intuición como justificación para creer (o no creer) en nada, así que ésta es una interpretación poco probable. Y si Mackie aceptara la validez de la intuición, entonces podríamos simplemente apelar a las diversas intuiciones citadas en este libro para demostrar que los valores objetivos existen. Dado que Mackie piensa que la cláusula de objetividad está integrada en todo el pensamiento y el discurso moral ordinario,[5] el caso de que el objetivismo sea contraintuitivo sería difícil de hacer.

Segunda: Quizá decir que *x* es raro sea decir que *x* es muy diferente de todas o de la mayoría de las cosas que existen. En este caso, la

premisa (2) se convierte en: "Si algo es muy distinto de otras cosas que existen, entonces (probablemente) no existe". ¿Por qué debemos aceptar este argumento? Consideremos esta lista parcial de lo que existe:

tiempo	estados mentales
espacio	estado físicos
números	propiedade estéticas
proposiciones	campos (por ejemplo, el campo gravitatorio)
sustancias	el pasado
propiedades	propiedades disposicionales
relaciones	propiedades morales

Se podría decir de cualquier elemento de esa lista que es totalmente diferente de todo lo demás que existe. Al menos eso parece, y sin embargo todas esas cosas parecen existir claramente. No se trata simplemente de que todo sea diferente de todo lo demás, sino de que cada una de las cosas mencionadas es muy diferente, diferente en especie, de todo lo demás. Así que, a primera vista, esta versión del argumento de la rareza no lleva a ninguna parte.

Tercera: Quizá el antirrealismo sobre la rareza sea correcto. Quizá la "rareza" sea la propiedad de provocar algún tipo de reacción en los observadores humanos, como sorpresa, asombro o incredulidad; quizá "*x* es raro" exprese una reacción no cognitiva hacia *x*; o quizá simplemente no exista tal cosa como la rareza. Si el antirrealismo moral es cierto, entonces creo que el antirrealismo sobre la rareza es casi con toda seguridad cierto. Pero entonces no está muy claro cómo llamar raro a algo puede considerarse una prueba en contra de su existencia.

Cuarta: Quizás decir que algo es raro es decir que no entendemos

claramente su naturaleza. Esta sería una versión de la tercera interpretación, y se aplica la misma objeción: no está claro cómo nuestra incapacidad para entender claramente algo es una prueba contra su existencia. Tampoco veo razones, en cualquier caso, para pensar que no comprendemos razonablemente bien las propiedades morales.

A este respecto, el antirrealista podría señalar el fracaso de los intuicionistas a la hora de explicar la naturaleza de las propiedades morales. Pero los intuicionistas han dicho muchas cosas sobre la naturaleza de las propiedades morales: que la bondad es una propiedad objetiva; que depende de las propiedades naturales de una cosa, pero que no es reducible a ellas; que tiene grados; que la conciencia de ella nos proporciona intrínsecamente razones para actuar; que es lo contrario de la maldad; que las cosas que la tienen son objetos apropiados de actitudes como la aprobación y la admiración; que es una propiedad que tienen en común cosas como la felicidad, el conocimiento, la benevolencia y la honestidad.

Sin embargo, los antirrealistas podrían señalar el fracaso de los intuicionistas a la hora de definir los conceptos morales. Pero esto no es razón para negar que entendemos los conceptos morales, ni para negar que se refieran a algo. A menudo entendemos un concepto lo suficientemente bien como para saber que se aplica a algunas cosas, sin poder definirlo. Yo no puedo definir "mesa", y no creo haber conocido a nadie que pudiera hacerlo.[6] Pero sigo sabiendo que hay mesas. Para un ejemplo más interesante, los filósofos analíticos han intentado durante 40 años definir el "conocimiento", pero ninguna definición ha sido generalmente aceptada y, en mi opinión, cada definición propuesta está sujeta a contraejemplos decisivos. Pero nadie cree que esto demuestre que no entendemos el "conocimiento" o que no se aplica a nada. De hecho, mi propia opinión es que nunca se ha producido una definición

correcta y no estipulativa de ningún concepto filosóficamente interesante. Obviamente, no puedo demostrarlo aquí, pero puedo informar de que no conozco ninguna definición de este tipo que haya sido generalmente aceptada en ninguna rama de la filosofía. La incapacidad de los intuicionistas para definir el "bien", por lo tanto, no muestra que la bondad sea extraña de una manera que justifique la negación de su existencia.

*8.2 Problemas con la superveniencia

*8.2.1 La objeción de Mackie

Todos o casi todos los realistas están de acuerdo en que las propiedades morales dependen de las propiedades no morales. Por ejemplo, si Pol Pot es malvado, debe haber algo más en él que lo haga malvado -por ejemplo, su desprecio por la vida humana, su personalidad tiránica, etc. Y sea cual sea esta propiedad o conjunto de propiedades, cualquier otra persona que la tuviera también sería mala. Expresamos esta idea diciendo que las propiedades morales "sobrevienen" a las propiedades no morales.[7]

J. L. Mackie considera misteriosa esta dependencia de las propiedades morales de las no morales:

> *¿Cuál es la conexión entre el hecho natural de que una acción es una pieza de crueldad deliberada -por ejemplo, causar dolor sólo por diversión- y el hecho moral de que es incorrecta? No puede ser una vinculación, una necesidad lógica o semántica. Sin embargo, no se trata simplemente de que las dos características se den juntas. La maldad debe ser de algún modo "consecuente" o "superveniente"; está mal porque es un acto de crueldad deliberada. Pero, ¿qué significa este "porque"?*[8]

Continúa preocupándose de que nuestra conciencia de esta relación requiera una misteriosa "intuición". Aparte de eso, la cita anterior es la exposición completa de la objeción.

Al igual que Mackie, no creo que el pasaje anterior articule ninguna objeción al realismo moral en absoluto. No veo nada ahí que proporcione una razón para negar que una propiedad pueda depender de otras propiedades. Mackie no ha identificado nada problemático sobre esta relación; todo lo que dice sobre ella es que no es una vinculación lógica o semántica. El resto del contenido del pasaje parece consistir en un tono de incredulidad.

Tal vez la idea de Mackie sea que esta relación de dependencia es extraña, como lo son las propiedades morales, o que la incapacidad de los intuicionistas para definirla es una razón para negar que exista. Si es así, se aplican mis respuestas de la sección 8.1.

Sin embargo, no niego que haya enigmas filosóficos sobre la superveniencia; veremos algunos en las siguientes subsecciones.

*8.2.2 Objeciones a la noción de rectitud *prima facie*

Algunos intuicionistas, entre los que me incluyo, creen que nuestras intuiciones nos dicen, no que ciertos tipos de acción son siempre correctos, sino que ciertos tipos de acción son correctos *prima facie*. Un tipo de acción correcta *prima facie* es un tipo de acción que tiende a ser correcta, o que es moralmente correcta en igualdad de condiciones.[9] La razón para introducir esta noción es que todas o la mayoría de las generalizaciones que se nos ocurren sobre lo que está bien o mal tienen excepciones. Por ejemplo, en general está mal quemar gatos. Pero si el gato tuviera una enfermedad mortal y contagiosa, y los microbios sólo se pudieran matar quemándolos, entonces estaría bien quemar un gato.

Además, resulta extremadamente difícil enunciar todas las condiciones en las que sería permisible quemar un gato. Por eso, algunos intuicionistas afirman que basta con ver que causar sufrimiento es malo prima facie: es malo mientras no existan circunstancias especiales que lo exculpen o justifiquen.

Peter Strawson objeta que esta idea entra en conflicto con la idea original de que las propiedades morales de una cosa se derivan de sus propiedades no morales. Si *B* se deduce de *A*, entonces *B* debe estar presente siempre que *A* esté, no sólo "tender a estar" presente cuando *A* está, o estar presente la mayor parte del tiempo cuando *A* está. Además, decir que un tipo de acción "tiende a ser mala" es simplemente decir que la mayoría de las acciones de ese tipo son malas. Pero una afirmación de la forma "La mayoría de las *A* son *B*" no puede ser una verdad necesaria. Sin embargo, los intuicionistas afirman que las verdades morales básicas son necesarias.[10]

Esta objeción tiene algunas premisas falsas. Los intuicionistas piensan que el estatus moral de una acción se desprende del conjunto de sus propiedades no morales, no de ninguna característica *prima facie* correcta o incorrecta. Tampoco "Quemar gatos es malo a primera vista" significa "La mayoría de las quemas de gatos son malas". Significa que el hecho de que un acto sea un caso de quema de gatos cuenta en contra de esa acción; es lógicamente compatible con esto que la mayoría de las veces existan circunstancias exculpatorias. (Imaginemos que la mayoría de los gatos del mundo tuvieran una enfermedad mortal, etc.)[11]

Para ver que esta concepción de lo correcto prima facie es perfectamente coherente, comparemos un ejemplo de la física clásica. Newton nos dice que un cuerpo acelerará en la dirección de la fuerza neta que actúe sobre él. También nos da una ley de la gravedad que rige una de las fuerzas que actúan sobre los cuerpos. Es físicamente necesario

que un cuerpo masivo experimente una atracción gravitatoria -algo que tiende a acelerarlo- hacia cualquier otro cuerpo masivo. Si no hay otras fuerzas, el cuerpo acelerará en esa dirección. Pero puede haber otras fuerzas, como las eléctricas o magnéticas, que superen esa atracción.

Prima facie, lo correcto y lo incorrecto son análogos a las "fuerzas" que empujan una acción hacia lo correcto o lo incorrecto.[12] Del mismo modo que, en la física clásica, la aceleración de cualquier cuerpo viene determinada por la configuración física del sistema al que pertenece, la corrección o incorrección de las acciones viene determinada por sus características no morales (incluidas las relacionales). Del mismo modo que es físicamente necesario que una determinada configuración genere una cierta cantidad de fuerza gravitatoria, es (metafísicamente) necesario que una determinada propiedad no moral genere un cierto grado de incorrección *prima facie*. Del mismo modo que la ley de la gravedad no significa que la mayoría de los objetos caigan (tal vez la mayoría sean sostenidos por fuerzas de contrapeso), el principio de la incorrección *prima facie* de la quema de gatos no significa que la mayoría de las quemas de gatos sean incorrectas.

Hay que reconocer que esta analogía es engañosa al menos en dos sentidos. En primer lugar, en el caso de la física clásica, existe una ley matemática simple para la composición de fuerzas (es decir, cómo calcular la "fuerza total" cuando se dan las fuerzas individuales): la ley del paralelogramo. No pretendo sugerir que la ley de la composición de lo correcto y lo incorrecto *prima facie* (el principio que nos lleva de las diversas características correctas e incorrectas de una acción a su grado general de corrección o incorrección) sea similar en su forma a la ley del paralelogramo para la composición de fuerzas, ni tampoco que tenga una formulación matemática sencilla. La única cuestión es que no hay nada incoherente en el marco general de los deberes *prima facie*.

En segundo lugar, en la física clásica, una fuerza puede ser contrarrestada por otra mayor, pero no puede ser eliminada cuando se dan las condiciones especificadas en la ley de fuerza correspondiente. Por ejemplo, la fuerza gravitatoria puede contrarrestarse, pero no hay forma de eliminar la fuerza gravitatoria que experimenta un cuerpo, siempre que éste se encuentre en las proximidades de otro cuerpo masivo. En cambio, las consideraciones morales no sólo pueden contrarrestarse, sino también eliminarse. Por ejemplo, supongamos que, *prima facie*, es correcto hacer feliz a la gente. Pero supongamos ahora que algún acto concreto produjera felicidad a un sádico al permitirle el placer de ver sufrir a otras personas. En este caso, la cuestión no es simplemente que el valor del placer del sádico se vea superado por el sufrimiento de sus víctimas, sino que la naturaleza sádica del disfrute que se produce le priva de todo valor moral. El placer del sádico no es en absoluto una razón a favor de la acción propuesta.[13]

Ahora bien, se podría pensar que esto crea un problema para la propia noción de derecho prima facie. Supongamos que un tipo de acción *prima facie* correcta se entiende como un tipo de acción tal que el hecho de que un acto sea de ese tipo cuenta como una razón a su favor. El tipo de ejemplo que acabamos de discutir muestra que es difícil encontrar un tipo de acción tal que el hecho de que un acto sea de ese tipo sea siempre una razón a su favor -siempre puede haber, al parecer, otras Circunstancias que neutralicen cualquier posible razón a favor de una acción.[14]

Hay al menos dos maneras de responder a esto. Por un lado, podríamos decir que este tipo de ejemplo sólo muestra que uno ha identificado mal lo que es *prima facie* correcto: no es la producción de felicidad lo que es *prima facie* correcto, podríamos decir, sino la producción de placeres inocentes. Alternativamente, podríamos

simplemente liberalizar la noción de lo correcto *prima facie*: podríamos decir que cuando un tipo de acción es correcta *prima facie*, entonces tenemos una razón para realizar actos de ese tipo en igualdad de condiciones, o en condiciones normales; esto es compatible con la posibilidad de circunstancias especiales que neutralizarían esta razón. Por supuesto, no hemos elaborado aquí una teoría moral satisfactoria; para ello, tendríamos que decir mucho más, sobre cuestiones como qué cosas son correctas *prima facie* y qué tipo de consideraciones neutralizan qué tipo de razones. Todo lo que pretendo aquí es defender el intuicionismo contra la acusación de que no dispone de una noción coherente de cómo se relacionan las propiedades morales con las propiedades naturales.[15]

Consideremos una segunda analogía. Nadie duda de que existen hechos objetivos sobre lo que es una buena jugada en el ajedrez, no moralmente buena, sino buena desde el punto de vista de ganar y no perder la partida. Que una jugada sea buena depende de la disposición de las piezas en el tablero (quizá junto con las capacidades de los jugadores). Hay reglas que ayudan a elegir buenas jugadas: intenta desarrollar tus piezas pronto, intenta controlar el centro del tablero, las torres son más valiosas que los caballos, etc. Todas estas reglas son derrotables; algunas posiciones posibles exigen violarlas. Sin embargo, estos principios pueden considerarse racionalmente correctos. Una persona sin experiencia en el ajedrez pero que conozca las reglas y el objetivo del juego podría ver que estas cosas tienen sentido.[16] Además, son verdades necesarias. (Es contingente cuáles son las reglas del juego, pero dadas estas reglas, es necesario, de manera similar a las verdades matemáticas, que estas reglas representen estrategias generalmente buenas.[17])

Esta analogía refuta el argumento de Strawson de que no puede ser

necesario que alguna clase de cosas tienda a tener una misma propiedad. Y demuestra la coherencia de la concepción intuicionista de la relación entre hechos evaluativos y naturales. No demuestra que esa concepción sea correcta -quizá la moral no sea como el ajedrez-, sino que demuestra que la concepción es coherente, al mostrar que existe realmente un ámbito en el que los hechos y nuestro razonamiento sobre ellos ejemplifican la estructura que el intuicionista atribuye a los hechos morales y al razonamiento moral.

*8.2.3 La objeción de Blackburn

Sea *A* un acto moralmente incorrecto. Sea *N* el conjunto completo de propiedades naturales de *A*.[18] Según Blackburn, la proposición (2) siguiente es analíticamente verdadera, pero (1) no lo es:

1. Si *A* tiene *N*, es incorrecto.
2. Si *A* tiene *N* y es incorrecta, entonces toda acción que tenga N es incorrecta.

En otras palabras, se puede sostener coherentemente (sin demostrar con ello que no se entiende el concepto de maldad), "*A* tiene *N* pero no está mal". Pero no se puede sostener coherentemente, "A tiene *N* y es incorrecto, pero otras cosas con exactamente las mismas propiedades naturales no son incorrectas". Blackburn piensa que el realista no puede explicar esto. Si la incorrección es una única propiedad extra de *A*, que no se deduce de las propiedades naturales de *A*, entonces ¿por qué no debería ser posible que haya un objeto que tenga incorrección, y otro objeto con las mismas propiedades naturales pero sin incorrección?[19]

Que (1) no es analítico se deduce de mi argumento anterior contra el naturalismo analítico (sección 4.2), así que lo concederé. (2) se deduce del principio de superveniencia de las propiedades morales. Pero no es

en absoluto obvio, como supone Blackburn, que (2) sea una verdad necesaria analítica y no sintética. Para que (2) sea analítica, debe haber una forma de sustituir expresiones sinónimas en (2) que dé como resultado una verdad lógica. Blackburn no hace ningún intento de demostrar esto, ni yo veo ninguna manera de hacerlo. Aun así, concedamos por el bien del argumento que (2) es analítica.

¿Cuál cree Blackburn que es la explicación correcta y antirrealista de la analiticidad de (2)? Piensa que si alguien rechazara la superveniencia, "traicionaría todo el propósito por el que moralizamos, que es elegir, elogiar, clasificar, aprobar o prohibir cosas sobre la base de sus propiedades naturales".[20] Si suprimiéramos "sobre la base de sus propiedades naturales" de esa frase, Blackburn no tendría explicación alguna. ¿Por qué debemos aprobar (etc.) las cosas sobre la base de sus propiedades naturales? ¿Por qué, desde un punto de vista antirrealista, sería inapropiado valorar las cosas por capricho, alabando una acción pero, sin razón alguna, condenando otra acción indistinguible? Blackburn sólo propone como un hecho bruto que tales actitudes no contarían como morales.

¿Por qué, entonces, no puede el realista moral proponer el siguiente hecho bruto: el concepto de propiedades morales incluye que dependen de propiedades naturales? Incluso si esto fuera todo lo que el realista tuviera que decir, su relato no sería menos explicativo que el de Blackburn.

Pero el realista dispone de otra explicación. Introduzcamos el término "hacer el bien" como un predicado que puede aplicarse, no a objetos o sucesos concretos, sino a características de objetos o sucesos. Los rasgos que hacen el bien son aquellos que tienden a hacer que la cosa que los tiene sea buena (esto está obviamente relacionado con la noción de corrección prima facie discutida anteriormente). Por ejemplo, podemos

decir: "Lo agradable es bueno". Análogamente, podemos decir: "Ser-un-acto-de-crueldad es hacer el mal". He aquí una hipótesis: (a) Nuestras intuiciones morales más fundamentales se refieren a hechos de la forma "*F* es hacer el bien", "*G* es hacer el mal", etcétera. Cuando una persona dice que, intuitivamente, matar está mal, en realidad está informando de la intuición de que ser-un-matador es hacer mal. (b) Decir que una cosa es buena es decir algo así como: sus rasgos que hacen el bien predominan sobre sus rasgos que hacen el mal, o que su conjunto total de rasgos naturales hace el bien.[21] Ésta parece una hipótesis plausible, consistente con la evidencia introspectiva. Si es cierta, entonces es analítico que si una cosa es buena, toda cosa cualitativamente indistinguible también lo es. Analíticamente, la única manera de que una cosa sea buena es que su conjunto total de propiedades naturales sea bueno. De ello se deduce que cualquier otro objeto que tenga ese conjunto total de propiedades tiene un conjunto total de propiedades que lo hacen bueno y, por tanto, es bueno. Esto da a Blackburn lo que pedía: una explicación de por qué la superveniencia de las propiedades morales es analítica que no implica que las normas morales individuales sean analíticas. Al realista le va ahora mejor resolviendo el enigma de Blackburn que al propio Blackburn.

*8.2.4 La objeción de Jackson

Frank Jackson sostiene que la superveniencia de lo moral sobre lo no moral es difícil de conciliar con la irreductibilidad de lo moral. Supongamos que la bondad se superpone a las propiedades naturales. Puede haber muchas maneras posibles de que una cosa sea buena, es decir, muchas propiedades naturales que harían que una cosa fuera buena. Sean estas propiedades naturales N_1, N_2, y así sucesivamente. Para cada una de ellas, es necesario que cualquier cosa que las tenga sea buena (ése es el principio de superveniencia). Puesto que estamos estipulando que

{N_1, N_2, ... } es una lista completa de las formas posibles de que una cosa sea buena, también es necesario que cualquier cosa que sea buena tenga al menos una de estas propiedades. Por lo tanto, hay dos predicados, "bueno" y "N_1 o N_2 o ...", que se aplican exactamente a las mismas cosas en cada situación posible.[22] Si estos predicados se refieren a la misma propiedad, esto no es problemático. Pero el intuicionista sostiene que aquí hay dos propiedades diferentes, una natural y otra no natural, y que "bueno" se refiere a la no natural. La objeción de Jackson aquí es sutil: a qué se refiere una palabra viene determinado normalmente por su uso -a qué tipo de cosas la aplicamos, incluyendo en qué situaciones hipotéticas la aplicaríamos. Pero "bueno" y "N_1 o N_2 o ..." tienen el mismo uso en ese sentido. Entonces, aunque hubiera dos propiedades, ¿qué haría que nuestra palabra "bueno" se refiriera a la propiedad no natural en lugar de a la natural? Además, aunque hubiera algo que hiciera que la palabra se refiriera a la propiedad no natural, ¿cómo aprenderían los niños que aprenden el idioma que "bueno" se refiere a la propiedad no natural en lugar de a la natural?[23]

Creo que la respuesta es bastante sencilla para un intuicionista de pura cepa. A qué se refiere una palabra no sólo lo determina aquello a lo que la aplicamos, sino también, al menos en parte, las intenciones de los hablantes. Cuando pensamos en ciertos tipos de acontecimientos (por ejemplo), vemos intuitivamente que tienen una propiedad evaluativa adicional, además de sus propiedades naturales, y nuestra intención es que la palabra "bueno" se refiera a esa propiedad. Pretendemos referirnos a la propiedad superveniente, no a la subveniente; a la propiedad cuyo reconocimiento constituye intrínsecamente una razón para la acción; a la simple propiedad que explica lo que N_1, N_2, etc. tienen en común. Además, cuando los niños aprenden la lengua por primera vez, consideran que "bueno" se refiere a la bondad y no a (N_1 o N_2 o ...), porque asumimos por defecto que una palabra desconocida se

refiere a la propiedad más simple y unificada (en oposición a disyuntiva) coherente con su uso.

La objeción de Jackson depende en realidad de otra suposición: "Nuestras mentes no albergan detectores de las propiedades no naturales en las que creen los mooreanos". Dada esta suposición, Jackson infiere correctamente que "la historia sobre la adquisición del lenguaje ético debe ser una historia sobre la captación de patrones en lo natural".[24] Dado esto, es realmente desconcertante cómo "bueno" podría referirse a otra cosa que no sea un patrón de propiedades naturales. Pero la suposición de Jackson obviamente plantea la cuestión contra el intuicionismo, y tampoco da ningún argumento contra la existencia de la intuición ética. Su argumento, por lo tanto, sólo tendría fuerza contra alguien que postulara propiedades de valor irreductibles pero se negara a admitir la intuición o cualquier otra facultad similar para conocerlas.

*8.2.5 Una analogía: la superveniencia de las propiedades lógicas

Las propiedades lógicas de un argumento son supervenientes a los significados y a la disposición de los símbolos que contiene. Esto es en algunos aspectos análogo a la superveniencia de las propiedades valorativas sobre las propiedades no valorativas. Al igual que es imposible que dos acciones difieran en cuanto a la corrección sin diferir en alguna propiedad natural, es imposible que dos argumentos difieran en cuanto a la validez sin diferir en el significado o la disposición de sus símbolos. Así como las propiedades éticas de una acción están determinadas por sus propiedades naturales, las propiedades lógicas de un argumento están determinadas por los significados y la disposición de sus símbolos (especialmente los símbolos lógicos). Y así como (muchos filósofos creen) que no hay un algoritmo general para determinar lo que

es moralmente correcto, no hay un algoritmo general para calcular lo que es lógicamente necesario. Para la última afirmación, me baso en el Teorema de Godel, suponiendo que deseamos contar las verdades de la aritmética como lógicamente necesarias.[25] (Si usted no considera las verdades de la aritmética como lógicamente necesarias, entonces sustituya aquí alguna noción de necesidad matemática, y cambie la analogía por una de la superveniencia de la necesidad matemática sobre los significados). Esto muestra que estos aspectos de la doctrina de la superveniencia de lo ético son perfectamente coherentes.

Una desanalogía entre la ética y la lógica deductiva se refiere a la noción de corrección *prima facie*. No existe una noción paralela de "validez *prima facie*"; las reglas de la lógica deductiva nos dicen que las inferencias de cierto tipo son siempre válidas, no que tienden a ser válidas. Sin embargo, podemos encontrar una analogía con la corrección prima facie si nos fijamos en el razonamiento no demostrativo.[26] Es posible que *A* apoye no demostrativamente a *B* pero que, si se añadiera cierta información adicional, *C*, a las creencias previas, *A* ya no apoyaría a *B*. Podríamos decir entonces que *A* es una prueba "*prima facie*" de *B*, o que es una prueba de *B* en condiciones normales. Por ejemplo, en relación con un sistema normal de creencias previas, el hecho de que Jim declare que posee una motocicleta es una buena prueba de que posee una motocicleta. Pero si también se sabe que Jim es un mentiroso patológico que lleva años haciéndose pasar falsamente por miembro de una banda de moteros, entonces el testimonio de Jim ya no es una buena prueba de que posea una moto. Así pues, parece que la lógica no deductiva es análoga a la ética en los aspectos relevantes. Sin embargo, nadie sugiere que estas características de la lógica no deductiva -el hecho de que sus relaciones de apoyo probatorio sean supervenientes y que sean derrotables- apoyen el antirrealismo sobre el apoyo probatorio.

8.3 ¿Cómo podemos entender "bueno"?

¿Cómo aprendemos inicialmente el significado de la palabra "bueno"?[27] Parece que hay dos maneras de explicar a alguien el significado de una palabra. La primera es dar una definición. Pero los intuicionistas niegan que "bueno" pueda definirse, así que eso queda descartado. La segunda forma es especificar algunas cosas a las que se aplica la palabra. Por ejemplo, se señala una serie de perros, diciendo "perro", y el niño se hace a la idea de que la palabra se refiere a cosas de ese tipo. No se puede señalar la bondad, ya que la bondad no es perceptible por los sentidos, pero tal vez se le pueden decir al niño algunos ejemplos de cosas que son buenas, con lo que se hace la idea de que "bueno" se refiere a lo que esas cosas tienen en común. El problema es que la gente no está de acuerdo con lo que es bueno, así que ¿cómo puede establecerse que todos entienden lo mismo por "bueno"?

Las sugerencias del párrafo anterior sobre cómo se aprenden las palabras están demasiado simplificadas. Casi nada del vasto vocabulario de una persona normal se adquiere porque alguien le enseñe expresamente las palabras, ya sea definiéndolas o dándole listas de ejemplos. Casi todas las palabras se aprenden oyendo su uso en conversaciones en las que no se hace ningún esfuerzo explícito por explicar la palabra.

En cualquier caso, hay un acuerdo sustancial sobre qué cosas son buenas, a nivel cotidiano. Casi todo el mundo califica de "buenas" la felicidad, la salud, la amistad, el conocimiento y la larga vida. Además, la gente tiende a perseguir las cosas que llama "buenas", intenta que otras personas las persigan y dice cosas que implican que la "bondad" de una cosa es una razón para perseguirla. Los niños pueden aprender que

"bueno" se refiere a la bondad porque esa interpretación da sentido a la mayoría de las cosas que se dicen con la palabra, y ninguna otra interpretación, comparablemente sencilla, se acerca a ello.

Una objeción relacionada es que "bueno" no puede referirse a ninguna propiedad, ya que se llaman "buenas" muchas cosas totalmente diferentes: personas, acciones, estados de cosas, hechos, acontecimientos. Se trata de un error: cosas muy distintas pueden tener alguna característica abstracta en común. Compárese con el hecho de que los números pueden aplicarse a objetos físicos, ideas, proposiciones, acciones, etc.: "tres" significa lo mismo en "tengo tres pares de zapatos" y "hay tres razones principales para comprar un PC". Lo mismo ocurre con las propiedades o relaciones lógicas, como la identidad y la identidad de conjuntos, o las "propiedades" metafísicas, como la existencia o la determinación. Asimismo, no hay ningún obstáculo para que la "bondad" se aplique a cosas muy dispares.

8.4 ¿Es el intuicionismo demasiado subjetivo?

Algunos han acusado al intuicionismo de convertirse en una forma de subjetivismo. Algunos se quejan de que los intuicionistas no proporcionan mucho método para llegar a conclusiones morales, o que no proporcionan ningún método para distinguir las intuiciones correctas de las incorrectas.[28] Esta última acusación se ha discutido anteriormente (sección 6.4). En cuanto a la ausencia de método, compárese el "método" por el que llegamos a muchas creencias sobre objetos físicos: percibimos las cosas, y luego creemos que las cosas son como las percibimos. Este es un "método" muy simple, pero que no impugna la objetividad del conocimiento perceptivo; no más, entonces, el "método" intuicionista comparativamente simple de llegar al conocimiento moral impugna la

objetividad de este último.

Otra versión de la queja es que el intuicionismo hace que la verdad de una teoría moral dependa de su conformidad con nuestras intuiciones y, por tanto, es una fuente de subjetivismo.[29] Esto es simplemente un malentendido de la teoría. La intuición es nuestro medio para conocer las verdades morales. No crea la verdad moral, como tampoco la percepción crea verdades sobre el mundo físico.

Aun así, podría pensarse que el intuicionismo defiende una epistemología preocupantemente subjetiva, en la medida en que las intuiciones -meros estados mentales internos- se convierten en la fuente última de justificación de todas las afirmaciones en un campo supuestamente objetivo. También esta objeción puede basarse en un malentendido: el intuicionismo, tal y como yo lo entiendo, no aboga por apelar a las intuiciones como prueba a partir de la cual inferir conclusiones normativas (véase más arriba, sección 5.6). Sin embargo, sí sostiene que el hecho de que una creencia moral esté justificada depende de los estados mentales del creyente, incluidas sus intuiciones. Una consecuencia de esto es que la justificación depende del observador: una persona puede estar justificada para creer lo que otra no está justificada para creer. Esto, sin embargo, lejos de ser una forma objetable de "subjetivismo", es cierto para todo tipo de creencias.[30] La justificación de las creencias perceptivas depende de las experiencias perceptivas de cada uno. La justificación de las creencias sobre el pasado depende de los recuerdos. La justificación inferencial depende de las demás creencias. El hecho de que la justificación de las creencias morales dependa de los estados mentales del creyente no es sorprendente ni motivo de alarma.

8.5 ¿Mis argumentos demuestran demasiado?

Supongamos que mis argumentos contra el antirrealismo moral tienen éxito. ¿Demostrarían los argumentos del mismo tipo el realismo sobre la belleza? Tal vez. Algunos filósofos aceptan el realismo estético en cualquier caso; además, la "belleza" puede ser en sí misma un término evaluativo - "bello" puede implicar "bueno"- por lo que este resultado no sería terriblemente sorprendente. La cuestión es más preocupante: ¿Demostrarían estos argumentos también el realismo sobre lo sexy, lo divertido y lo genial (como en "Matrix es una película genial")? El realismo sobre tales propiedades es considerablemente menos atractivo que el realismo sobre el valor.[31]

Para empezar, consideremos el no cognitivismo. Dado que frases como "Creo que Steven es gracioso" y "Si Steven es gracioso, el público se reirá" tienen sentido y a veces se afirman adecuadamente, parece que el no cognitivismo sobre la gracia se enfrenta más o menos a los mismos problemas que el no cognitivismo sobre el valor. Los argumentos lingüísticos probablemente funcionarán exactamente de la misma manera, con casi la misma verosimilitud, en el caso de cualquier oración que adopte la forma gramatical de aserción, ya que cualquiera de esas oraciones probablemente podrá incrustarse en oraciones más amplias, al igual que las afirmaciones evaluativas. Este argumento no requiere ninguna intuición específica sobre la naturaleza de la diversión.

El argumento contra el "nihilismo" de lo divertido también parece análogo al argumento contra el nihilismo de los valores: el argumento consistiría en mostrar algunas cosas paradigmáticamente divertidas, cosas de las que casi todo el mundo se reiría. Puede que haya que buscar un poco, ya que los sentidos del humor varían, pero supongamos que encontramos tal paradigma.[32] Supongamos que se trata de este chiste:

> Un ingeniero, un científico, un matemático y un filósofo están juntos en Irlanda, cuando ven una oveja negra sola en un campo. El ingeniero dice: "Oye, parece que las ovejas de por aquí son negras". El científico le mira con escepticismo y responde: "Bueno, al menos algunas lo son". Luego, el matemático les echa otra mirada escéptica y dice: "Bueno, al menos una de ellas lo es". Por último, el filósofo se dirige a ellos y dice: "Bueno, al menos en un lado".

Imagina que al oír eso, todos se ríen a carcajadas (incluido tú). Te invito a que estés de acuerdo en que ha sido gracioso. ¡Qué absurdo sería para el nihilista afirmar, a causa de una dudosa teoría filosófica, que el chiste no era divertido!

Aun así, podríamos considerar que el nihilismo de la diversión es menos problemático que el nihilismo del valor. Renunciar a las afirmaciones de diversión implicaría un cambio menor en nuestros sistemas de creencias. No tendríamos que dejar de reírnos de las cosas, ya que la risa no es una creencia, ni está causada o "justificada" por creencias sobre la diversión.

Prefiero no negar que el chiste anterior sea gracioso. A mí me parece gracioso. Tal vez nuestra mejor apuesta sea una teoría subjetivista de la diversión. Repasemos los problemas del subjetivismo del valor del capítulo 3 y veamos si se aplican a la subjetividad del humor:

i) *El problema de las actitudes objetivas*: Imaginemos a alguien cuyo sentido del humor difiere radicalmente del nuestro (pero no de una manera que pueda ser moralmente ofensiva, como reírse del sufrimiento de otro). Sale a la calle y ve un palo cualquiera, con lo que rompe a reírse histéricamente. El subjetivista dice: "Para él, el palo es divertido". Esto no parece obviamente incorrecto, del mismo modo que "Para Ted Bundy, el asesinato es permisible" es obviamente incorrecto. Por otro lado, algunos probablemente

describirían a esta persona de la siguiente manera: "Se ríe de cosas que no son divertidas". Quizás en este contexto, "es gracioso" significa aproximadamente, "tendería a hacer reír a la gente típica". Esto implica que, si la mayoría de nosotros estamos dispuestos a reírnos de los palos, entonces los palos son graciosos. Una vez más, esto no parece inaceptable de la forma en que lo es "Si la mayoría de nosotros aprueba la esclavitud, entonces la esclavitud es buena".

ii) *El problema del desacuerdo*: Después de oír el mismo chiste, una persona dice: "Me ha hecho gracia", y otra dice: "No me ha hecho gracia". El subjetivista dice que no están en desacuerdo: uno de ellos sólo quiere decir que el chiste le ha hecho gracia, mientras que el otro dice que no le ha hecho gracia. Esto no es inverosímil. La frase "El chiste me hace gracia" tiene sentido, y la primera persona puede estar dispuesta a aceptar que eso es lo que quería decir. Es muy poco probable que insista en que la segunda persona está equivocada, ni que presente argumentos para demostrar que el chiste era gracioso, a no ser, quizá, que diga que la mayoría de la gente se reiría con él. En cambio, la frase "Para mí el aborto está mal" no tiene sentido y no puede sustituirse por "El aborto está mal". Si alguien dice "El aborto no es malo", los antiabortistas dirán que se equivoca y darán argumentos para demostrar que el aborto es malo. Sus argumentos no apelarán a la afirmación de que la mayoría de la gente desaprueba el aborto.

iii) *El problema de la arbitrariedad*: Argumentamos que el subjetivismo valorativo hace que las actitudes o los mandatos de la persona o el grupo en cuestión sean arbitrarios. Preguntas como "¿Por qué deberíamos obedecer reglas arbitrarias?" y "¿Cómo podría ser bueno que alguien apruebe *x* sin razón?" parecen entonces apropiadas. El subjetivismo del humor, igualmente, hace que nuestra risa sea arbitraria, no en el sentido de que no pueda ser explicada causalmente, sino en el sentido de que no puede ser

justificada. Pero esto no es en absoluto perturbador, del modo en que lo sería la arbitrariedad en la ética. La pregunta "¿Por qué deberíamos reírnos de ese chiste?" no se plantea de forma natural, y "¿Cómo puede ser que el hecho de que estemos dispuestos a reírnos de *x* sin ninguna razón haga que *x* sea divertido?" parece poco convincente: generalmente no esperamos que la gente tenga razones (es decir, justificaciones) para reírse, en el sentido de que nosotros podríamos tener razones para nuestras acciones.

iv) *El problema de la falibilidad*: El falibilismo sobre el humor también parece poco convincente. Supongamos que alguien dice: "Ese chiste me hace gracia, pero podría estar equivocado". Me cuesta entenderlo, a no ser que el orador se refiera a lo que a la mayoría de la gente le haría gracia, o algo por el estilo. (Imagínese a alguien que dirige un club de comedia que intenta decidir a quién contratar para el entretenimiento de la noche. Le parece que Sally es graciosa, pero más tarde decide que se equivocó sobre su gracia cuando el público comienza a abuchear).

Parece, pues, que el subjetivismo del humor es mucho más defendible que el subjetivismo moral.

Un punto más. Muchos ejemplos de propiedades "obviamente subjetivas" pueden tener componentes evaluativos, y esto obstaculizará cualquier intento de hacer análisis puramente subjetivos. Si el asesinato de bebés impresionara a mucha gente, entonces el asesinato de bebés sería genial" es falso porque el asesinato de bebés es malo, y "genial" implica "bueno" (independientemente de lo que implique). Del mismo modo, en algunos casos, una persona puede argumentar que algo "no es divertido" por motivos morales; por ejemplo, que es inapropiado reírse de algo porque al hacerlo se ofende a los demás o se expresa una mala actitud. Tal vez "divertido" signifique "no es malo reírse de algo". "Sexy" puede implicar también "no es malo (o anormal) sentir atracción

sexual por algo". Si esto es así, entonces, por supuesto, estos conceptos se resistirán a los tratamientos puramente subjetivos, pero eso no supondría un reto para los argumentos de este libro.

8.6 Evolución y ética

8.6.1 La objeción evolutiva al realismo

Los Sociobiólogos (o "psicólogos evolutivos", en la jerga actual) han ofrecido explicaciones para el comportamiento moral y las creencias morales, en términos de la teoría de la evolución. Como la evolución nos lleva a esperar egoísmo, el principal problema es explicar por qué solemos ayudar a los demás a costa de nosotros mismos. Entre las explicaciones parciales propuestas figuran:

Selección de parentesco: Ayudamos a los miembros de nuestra propia familia, a menudo a costa de nosotros mismos, porque los miembros de la familia comparten muchos de los mismos genes. Ayudar a la propia familia es, por tanto, una forma indirecta de aumentar la reproducción de los propios genes.

Altruismo recíproco: ayudamos a los demás porque al hacerlo tendemos a que ellos quieran ayudarnos a su vez. Si se odia a los demás o se rompen los acuerdos con ellos, se tiende a enfadarlos y a hacer que nos odien.

Efectos de reputación: Nos comportamos amablemente porque así mejoramos nuestra reputación general en nuestra comunidad, haciendo que los demás nos consideren buenos socios comerciales, etcétera. Quien rompe acuerdos o ataca a la gente es probable que sea rechazado.[33]

Los sociobiólogos no dicen que pasemos por los tipos de razonamiento

anteriores antes de decidir qué hacer, sino que estos argumentos explican por qué nuestros sentimientos de querer ser amables con los demás tienen valor reproductivo.

Algunos dicen que la capacidad de la teoría evolutiva para explicar nuestras creencias morales socava el realismo moral, en particular el intuicionismo ético. Hay tres argumentos para ello.[34]

Primero: Puesto que la sociobiología proporciona explicaciones adecuadas para nuestras creencias morales, y estas explicaciones no apelan a valores objetivos, no hay necesidad de plantear valores objetivos. No nos ayudan a explicar nada, y nuestra teoría general es más sencilla sin ellos. Tampoco necesitamos plantear una facultad cognitiva especial de intuición moral, cuando podríamos explicar las creencias morales apelando a las emociones, que a su vez se explican evolutivamente.

Segundo: Incluso si existen valores objetivos, la explicación evolutiva de nuestras creencias morales no nos da ninguna razón para esperar que nuestras intuiciones tiendan a corresponder a la verdad moral. Esto contrasta con otras facultades cognitivas. Un organismo con una visión deficiente tendrá peores resultados en términos de supervivencia y reproducción que un organismo con una visión precisa. Una persona que confunde tigres con melocotoneros tendrá menos posibilidades de sobrevivir. Pero no tenemos constancia de cómo mejoraría la aptitud reproductiva si las creencias morales fueran correctas.

Tercero: No sólo carecemos de una explicación de por qué deberíamos tender a tener creencias morales correctas, sino que la explicación sociobiológica de las creencias morales proporciona motivos positivos para esperar que nuestras creencias morales no sean fiables,

suponiendo que exista una verdad moral objetiva. Esto se debe a que la evolución tendería a darnos fuertes prejuicios a favor de nosotros mismos, de nuestros parientes y de los miembros de nuestro propio grupo social. La evolución moldearía nuestras actitudes morales hacia sus propios propósitos, para asegurar un comportamiento que aumente la aptitud reproductiva, independientemente de lo que realmente sea moralmente correcto.

El primer argumento sugiere que no hay valores objetivos. Los dos últimos sugieren que, aunque los haya, no tenemos forma de conocerlos.

8.6.2 La carga del realista

Es importante recordar que los intuicionistas no dicen que tengamos un "sentido moral especial", es decir, una facultad separada dedicada únicamente a conocer las verdades morales. La intuición es una función de la razón. La intuición moral difiere de la intuición matemática del mismo modo que la percepción de los coches difiere de la percepción de los árboles, es decir, simplemente por tener objetos diferentes. Por lo tanto, no se requiere una explicación especial para explicar cómo evolucionó "la facultad de la intuición moral".

Consideremos un caso análogo: ¿por qué tenemos la capacidad de ver las estrellas? Al fin y al cabo, es de suponer que nuestros antepasados evolutivos lo habrían hecho igual de bien si sólo vieran cosas en la Tierra. Por supuesto, esta es una pregunta tonta. Podemos ver las estrellas porque tenemos visión, que es útil para ver cosas en la Tierra, y una vez que tienes visión, acabas viendo cualquier cosa que esté enviando luz en tu dirección, sea algo útil de ver o no. Del mismo modo, una vez que se tiene inteligencia, se acaba por aprehender el tipo de cosas que pueden ser conocidas por la razón, tanto si son útiles de conocer como si no. Así, los humanos son capaces de aprender a jugar al

ajedrez a un nivel increíblemente sofisticado, a pesar de que el entorno de los australopitecos no contenía tableros de ajedrez. Si algunas verdades evaluativas pueden conocerse por medio de la razón, es probable que entonces las conozcamos, tanto si son útiles para conocerlas como si no.[35]

Nótese que la teoría moral no necesita ser validada positivamente por la teoría de la evolución, como no lo necesita ninguna otra rama del conocimiento. Para saber que 2+2=4, no tenemos que consultar primero a los sociobiólogos y ver si la evolución predice que los humanos tendrían creencias matemáticas correctas.[36] Para saber que tengo un ordenador delante, no tengo que ver primero si la evolución predice que los humanos tendrían creencias físico-objetuales correctas. Esto es afortunado, ya que si esto fuera un requisito, no podríamos obtener ningún conocimiento del mundo físico antes de conocer la teoría de la evolución, en cuyo caso no tendríamos forma de verificar esa teoría para empezar. Hemos visto que algunas creencias morales pueden tomarse razonablemente como fundacionales, al igual que algunas creencias matemáticas y perceptivas. La justificación inicial de las creencias sobre los hechos morales no deriva de su encaje en una explicación sociobiológica, ni de la capacidad de los hechos morales para explicar los hechos naturales.

El antirrealista, por tanto, no debería argumentar contra el realismo basándose en que la sociobiología no predice que tendríamos conocimiento moral. Ese argumento no sería más plausible que el argumento de que deberíamos rechazar la aritmética porque la sociobiología no predice que tendríamos conocimiento de ella. Pero el antirrealista puede argumentar que la sociobiología predice que no tendríamos conocimiento moral. Aquí, el supuesto no sería que la moralidad necesita validación externa, sino sólo que puede ser socavada

por su tensión con otras creencias justificadas. Por lo tanto, en mis siguientes respuestas, sólo necesito argumentar que la teoría de la evolución no predice de forma determinante la ausencia de conocimiento moral.

8.6.3 ¿Qué tan buenas son las explicaciones evolutivas de la ética?

Las explicaciones evolutivas de la ética no son impresionantes, ciertamente no lo suficiente como para que su existencia justifique hacer una revisión radical de nuestro sistema de creencias y rechazar todas las creencias morales.

¿Qué sería un conjunto de explicaciones "impresionantes"? En primer lugar, sería uno que prometiera explicaciones específicas para todas las principales intuiciones éticas, incluidas las que son los ejemplos más plausibles de principios racionales. Debería explicar intuiciones como la superveniencia del valor, la universalidad, la asimetría y la transitoriedad del "mejor que", los derechos y el imperativo categórico. Debería prometernos explicaciones para los detalles de nuestras intuiciones sobre situaciones particulares, como las intuiciones aparentemente divergentes en el problema del trenvía y el ejemplo de la recolección de órganos.[37] Con pocas excepciones, las explicaciones evolutivas de la ética pretenden explicar el comportamiento moral sólo en sus más vagas líneas, considerando que la moralidad consiste en poco más específico que el imperativo "cooperar". Las únicas creencias morales relativamente específicas que los sociobiólogos han explicado hasta ahora son las que inicialmente parecen más adecuadas para la explicación evolutiva y las peores candidatas a creencias morales racionales, por ejemplo, las creencias morales sobre el incesto, las obligaciones especiales para con la familia y las prácticas

reproductivas.[38]

En segundo lugar, se trataría de un conjunto de explicaciones que implican predicciones únicas, o al menos bastante específicas, sobre el contenido de la moral. Es decir, debería ser evidente que el marco explicativo utilizado no podría dar cabida a una amplia gama de creencias y comportamientos morales diferentes, sino que requiere que las creencias morales adopten formas bastante específicas, de modo que la hipótesis de que las creencias morales son adaptaciones evolucionadas habría sido falsificada si hubiéramos tenido intuiciones morales muy diferentes de las que tenemos en realidad. Por ejemplo, nos gustaría tener un marco explicativo que pudiera explicar por qué la Regla de Oro sería aceptada, pero no podría dar cabida a la aceptación generalizada del egoísmo (dado que el egoísmo es contrario a nuestras intuiciones morales).

¿Por qué imponer este requisito? Michael Ruse argumenta que la disponibilidad de explicaciones evolutivas de la ética es una prueba convincente contra la realista moral. Pero si las explicaciones evolutivas de la ética hubieran estado disponibles con toda probabilidad independientemente de que el realismo fuera cierto -por ejemplo, porque se puede pensar en una explicación evolutiva para prácticamente cualquier cosa-, entonces el hecho de que estén disponibles no es una prueba significativa contra el realismo.

Muchos se han quejado de que la sociobiología es "infalsificable".[39] Mi afirmación es más débil: No estoy criticando la sociobiología en general, ni estoy afirmando que la idea de una explicación evolutiva para las creencias éticas no pueda ser falsada por ninguna prueba. Mi afirmación es que hay una gama tan amplia de creencias éticas concebibles a las que se podría dar una explicación evolutiva, que el hecho de que nuestras creencias éticas reales sean susceptibles de una (si

lo son) no nos dice mucho. Si hubiéramos tenido la intuición generalizada de que el egoísmo ético era cierto (la acción correcta es siempre la que mejor sirve a los propios intereses), ¿no habrían reclamado los sociobiólogos una explicación bastante directa para ello? Pero nuestras intuiciones morales reales (en esta dimensión de egoísmo frente a altruismo) son casi lo contrario. Si los sociobiólogos pueden explicar esto también, entonces la gama de creencias morales que podrían acomodar en principio es muy amplia. Para dar un ejemplo más extremo de la flexibilidad del paradigma, algunos sociobiólogos han llegado a afirmar que la homosexualidad es una adaptación.[40] Si incluso la homosexualidad puede encajar en el paradigma sociobiológico, no es de extrañar que algunas creencias morales también puedan hacerlo.

8.6.4 Una explicación evolutiva de la percepción moral

En mi opinión, la explicación correcta de por qué tenemos intuición ética se referirá a lo que sea la explicación de por qué tenemos razón e inteligencia en general. Sin embargo, ahora quiero dejar de lado este punto y sugerir que existe una explicación evolutiva adicional para explicar por qué las personas tienen creencias éticas correctas. Esta explicación requiere cuatro supuestos:

a) Un sistema moral tiene en cuenta de forma justa los intereses de todos. Esto no es todo lo que hay en la moral, pero es un aspecto (quizás analíticamente) necesario de la moral. Un sistema de reglas que ignore los intereses, o que tenga en cuenta algunos intereses excluyendo arbitrariamente otros, no sería un sistema moral. Hay diferentes formas posibles de tener en cuenta los intereses de forma justa, por ejemplo, formas consecuencialistas y formas basadas en los derechos. No intentaré analizar aquí el término "justo", pero debe entenderse como un contraste con "parcial".

b) La cooperación pacífica proporciona beneficios a todos que generalmente superan cualquier beneficio que pudiera obtenerse mediante la explotación de los demás. El conflicto con otros causa daños que generalmente superan los beneficios que se podrían obtener sin cooperar. Supongamos que usted tiene 100 dólares que yo quiero. Decido ir a tu casa y cogerlos, lo que da lugar a una pelea. Como soy más fuerte, gano la pelea y me llevo los 100 dólares. Sin embargo, me rompo un brazo, un daño mucho peor que la pérdida de 100 dólares. Tú estás incluso peor que yo (con dos brazos rotos, digamos), pero eso es irrelevante; mi objetivo era beneficiarme a mí mismo, no perjudicarte a ti, así que he fracasado.

c) La cooperación pacífica requiere un conjunto de reglas acordadas. Si yo me rijo por unas normas que me asignan la propiedad de esos 100 dólares, mientras que tú te riges por unas normas que te asignan la propiedad de esos mismos 100 dólares, es probable que se produzca un conflicto. Si los dos tratamos de cooperar en alguna tarea, pero tenemos interpretaciones muy diferentes de lo que cada uno debe hacer, es probable que nuestra cooperación se rompa.

d) El acuerdo sobre un conjunto de normas sociales es improbable a menos que esas normas tengan en cuenta de forma justa los intereses de todos. Si me rijo por unas normas sesgadas a mi favor, es poco probable que encuentre a alguien que esté de acuerdo con ellas. Para cooperar pacíficamente con los demás, debo aceptar normas que tengan en cuenta mis intereses y los de los demás.

Esto no nos lleva todavía al conjunto correcto de reglas morales. Mi sugerencia, sin embargo, es que las personas con capacidad de intuición moral disfrutarían de la ventaja de acordar con los demás un conjunto de reglas que permita la cooperación social, ya que tiene en cuenta de forma justa los intereses de todos. Esta no es la única manera de lograr dicha

coordinación, pero es una de ellas. Las personas que carecen de una intuición moral precisa podrían no ver ninguna razón para aceptar un conjunto de reglas sociales, o podrían inclinarse por aceptar un conjunto de reglas sesgadas a su favor. Incluso si aceptaran un conjunto de normas que tuvieran en cuenta los intereses de todos, es probable que siguieran sin estar de acuerdo, ya que hay muchos conjuntos de normas posibles. Como sólo hay una verdad, los que tienen una percepción moral precisa evitan este problema.[41]

Mis afirmaciones sobre esta explicación son modestas. Tal vez sea más o menos correcta; tal vez no. En cualquier caso, demuestra que es posible que una percepción moral precisa tenga valor para la supervivencia.

No niego que la sociobiología pueda enseñarnos algo sobre la ética. Como se ha señalado, la evolución puede habernos dotado de prejuicios que afectan a nuestros juicios morales. La sociobiología puede ayudarnos a identificar estos sesgos, y así corregirlos, mejorando así nuestra cognición moral.

8.7 ¿Es el intuicionismo demasiado revisionista?

Dado que cabe esperar que las creencias y prácticas morales actuales se correspondan en gran medida con nuestras intuiciones morales, es natural suponer que el intuicionismo conduce a opiniones éticas muy conservadoras. Pero esto no tiene por qué ser así. Consideremos el caso análogo de nuestro conocimiento del mundo exterior. La ciencia nos dice que el sol es 1,3 millones de veces más grande que la Tierra, que los campos de fuerza invisibles llenan todo el espacio y que los objetos materiales ordinarios están compuestos por partículas diminutas e incoloras en rápido movimiento, con grandes espacios entre ellas. Estas

conclusiones son sorprendentes y se alejan bastante de la forma en que aparecen las cosas a la observación casual. Sin embargo, los científicos no poseen una nueva facultad cognitiva de la que la gente común está privada; la ciencia se basa, única o principalmente, en la observación.

Del mismo modo, el razonamiento moral puede llevar a conclusiones sorprendentes y revisables, a pesar de que este razonamiento debe basarse en última instancia en la intuición. Un intuicionista aceptará las intuiciones como fuente de justificación *prima facie*, pero no debería aceptar las intuiciones acríticamente, igual que un científico no acepta acríticamente las observaciones aparentes. Como vimos en el capítulo 6, hay muchas maneras en que nuestras intuiciones y creencias pueden equivocarse, y el intuicionista sabio estará en guardia contra estas fuentes de error.

Desde la perspectiva intuicionista, hay buenas razones para creer que es necesaria una teoría moral revisada. Muchas de las creencias morales más extendidas parecen más productos del interés propio, del adoctrinamiento cultural o de los imperativos de la selección natural que productos de una reflexión cuidadosa y racional.[42] Por ejemplo, la creencia generalizada de que la poligamia es mala se deriva aparentemente de nuestra cultura particular.[43] La mayoría de las culturas han practicado la poligamia sin considerarla mala en absoluto, y pocos de los que sostienen la creencia de que la poligamia es mala pueden ofrecer alguna razón plausible para ese juicio; tampoco es lo malo de la poligamia el tipo de cosa que esperaríamos que fuera simplemente una verdad evaluativa fundamental, incapaz de mayor explicación. Del mismo modo, pocos están dispuestos a dar una explicación racional de la supuesta ilicitud del incesto; la mayoría de nosotros, al parecer, simplemente pensamos que está mal.[44] Algunos han especulado que esto tiene una explicación evolutiva, relacionada con el mayor riesgo de

defectos congénitos derivados del incesto.[45] Y la creencia generalizada de que, mientras que matar a seres humanos para alimentarse es uno de los males más atroces imaginables, matar a miembros de cualquier otra especie para alimentarse no está ni siquiera un poco mal, parece explicarse más fácilmente por el interés propio, la cultura y/o la selección natural que por una perspicacia racional. De nuevo, la creencia apenas parece una verdad axiomática evaluativa y, sin embargo, la mayoría de los que la sostienen son incapaces de aducir ninguna razón creíble para ello.[46] Los antirrealistas toman todo esto como una prueba contra el realismo moral;[47] los intuicionistas, sin embargo, simplemente lo tomarán como una razón para rechazar las creencias morales particulares en cuestión.

Nos llevaría demasiado lejos discutir aquí los méritos de éstas y otras creencias morales comunes. Pero supongamos que la reflexión de los intuicionistas más cuidadosos y minuciosos les lleva a negar muchas, quizá la mayoría, de las creencias morales más comunes. Entonces podríamos preocuparnos de que el intuicionismo sea demasiado revisionista. La idea aproximada que tengo en mente es la siguiente: Los significados de (todas o la mayoría de) las palabras de nuestro idioma, incluidas "incorrecto" y "permisible", vienen determinados por cómo las utilizan habitualmente quienes hablan el idioma. Por esta razón, no puede darse el caso de que una palabra se utilice mal normalmente; un "mal uso" suficientemente común de un término simplemente genera un nuevo significado para el término. Por la misma razón, una persona que utiliza una palabra de forma radicalmente distinta (en el sentido de aplicarla a una clase muy diferente de cosas) a como se utiliza habitualmente puede ser acusada simplemente de malinterpretar la palabra. Imaginemos a alguien que niega que animales como los pastores alemanes, los caniches y los terriers escoceses sean perros, sino sillas. A esta persona se le acusaría con razón, sólo por este motivo, de

malinterpretar la palabra "silla". Esto no quiere decir que las personas que entienden el término "silla" a veces puedan confundir objetos que no son sillas con sillas. Pero una persona no puede entender la palabra inglesa "chair" (silla) y, sin embargo, aplicarla sistemáticamente a un tipo de objeto muy distinto de aquel al que la mayoría de los angloparlantes la aplican. Uno podría decir: no puede resultar que los pastores alemanes sean sillas, porque eso no es lo que entendemos por "silla".

Del mismo modo, me imagino a un antirrealista argumentando lo siguiente:

1. Si el intuicionismo es verdadero, entonces la incorrección existe y es una propiedad objetiva de la que somos conscientes por intuición racional. (Premisa.)

2. No hay ninguna propiedad objetiva de la que seamos conscientes por intuición racional, que tenga en común la mayoría de las cosas que la gente suele llamar "malas". (Premisa.)

3. Por lo tanto, si el intuicionismo es verdadero, entonces la maldad existe, pero la mayoría de las cosas que la gente suele llamar 'mal' no la tienen. (De 1, 2.)

4. No puede ser que la maldad exista pero que la mayoría de las cosas que la gente suele llamar "malas" no la tengan. (Premisa.)

5. Por lo tanto, el intuicionismo es falso. (De 3, 4.)

La premisa (1) simplemente reafirma (parte de) la teoría del intuicionismo ético. La premisa (2) se apoya en el hecho de que las personas -basándose en las emociones y otras causas no racionales- suelen juzgar "incorrectas" acciones que no parecen tener en común ninguna característica objetiva relevante. La premisa (4) se apoya en la concepción general del lenguaje que acabamos de analizar, según la cual

el significado de una palabra viene determinado normalmente por su uso.[48]

¿Cómo debería responder un intuicionista? El primer problema es que el argumento de la premisa (4) elude una importante distinción entre dos sentidos en los que una palabra puede ser "mal empleada". En el primer sentido, que es el más propio, una palabra está mal empleada cuando se utiliza para significar algo que no significa según las convenciones del lenguaje. Por ejemplo, en la frase "Al decir que no subiría los impuestos a la clase media, el Presidente dedujo que podría subir los impuestos a otros grupos", la palabra "dedujo" se utiliza incorrectamente con el significado de "implícito". En el segundo sentido, puede decirse que una palabra está "mal empleada" cuando, grosso modo, se aplica a algo que no tiene la propiedad a la que la palabra se refiere. Por ejemplo, en la frase "Elizabeth Proctor es una bruja", la palabra "bruja" se aplica incorrectamente a Proctor, que en realidad no es una bruja. Sin embargo, no está mal empleada en el primer sentido: no se utiliza para significar algo que la palabra no significa normalmente. Ahora bien, es cierto, por la razón expuesta anteriormente (que el uso determina el significado de las palabras), que una palabra no puede ser generalmente mal empleada en el primer sentido. Pero, como muestra el ejemplo de la palabra "bruja", una palabra puede usarse mal en el segundo sentido; de hecho, la palabra "bruja" se ha usado mal en este sentido cada vez que se ha aplicado a alguien, ya que, como supongo, no hay brujas. Del mismo modo, es concebible que la palabra "incorrecto" sea comúnmente "mal utilizada" en el sentido de ser aplicada a cosas que no tienen la propiedad de ser incorrectas, aunque no, por supuesto, en el sentido de ser utilizada para significar algo que normalmente no significa. Así pues, la premisa (4) parece falsa; en cualquier caso, no se sostiene por el mero hecho de que los significados de las palabras estén determinados generalmente por el uso.

Aun así, no creo que la acusación de que el intuicionismo es demasiado revisionista carezca por completo de plausibilidad. Aunque la objeción que hemos discutido no demuestra que sea conceptualmente incoherente atribuir un error moral masivo a las personas, puede ser cierto que el significado de "wrong" en inglés es tal que excluye la posibilidad del tipo de error que he hipotetizado. Tal vez la comprensión de la palabra por parte de la mayoría de la gente esté tan estrechamente ligada a supuestos paradigmas de maldad como el incesto, la poligamia y la limpieza del retrete con la bandera nacional[49] que es analíticamente cierto que, si existe la maldad, entonces esas acciones la tienen. Tal vez, al mismo tiempo, el concepto de "mal" de la mayoría de la gente está lo suficientemente ligado a la noción de las razones para la acción como para que el hecho (si es que es un hecho) de que existan suficientes razones, consideradas en su totalidad, para no comer animales no humanos no sea suficiente para establecer que esa comida es mala. Digo "quizás" porque, aunque no creo que esto sea así, no estoy seguro de que no lo sea. Y por eso quiero abordar la cuestión de cómo debería reaccionar un intuicionista si estas cosas son el caso y si, al mismo tiempo, la reflexión no revela ninguna propiedad objetiva relevante que tengan en común el incesto, la poligamia y la limpieza del inodoro con la bandera.

Brevemente, creo que un intuicionista podría, sin que su posición pierda interés, aceptar la acusación de usar algunas palabras, como "moralmente correcto" y "moralmente incorrecto", en sentidos no estándar. Pretendo que el uso que hago de la palabra "incorrecto" se refiera, a grandes rasgos, a la propiedad sobre la que (la mayoría de) lo siguiente es cierto: es objetiva; no es una propiedad física; sólo las acciones intencionadas pueden tenerla; el asesinato, el robo, la ruptura de promesas y otras acciones que dañan a otros o violan los derechos de los demás suelen tenerla; cuando una acción la tiene, hay suficientes

razones consideradas en su conjunto para no llevarla a cabo; la culpa y el remordimiento son reacciones típicas y apropiadas ante la conciencia de que una acción propia la tiene; la culpa es una respuesta típica y apropiada ante la conciencia de que una acción ajena la tiene; una acción sólo la tiene si el agente podría haberse abstenido de realizarla; una acción no puede tenerla si la acción es permisible, elogiable u obligatoria. Cuando digo que la poligamia no es mala, quiero decir que la poligamia no tiene ninguna característica en la que todas o la mayoría de esas cosas sean ciertas.

La más importante de las condiciones es que las acciones erróneas son acciones para las que hay razones suficientes y objetivas para no realizarlas. El núcleo de mi intuicionismo es la afirmación de que a veces podemos ver, intuitivamente, razones objetivas para actuar, razones que no dependen de ningún deseo contingente ni de otras respuestas psicológicas nuestras. Si resulta, como podría ser el caso, que un número suficientemente grande de angloparlantes no utilizan la palabra "wrong" ni siquiera de esta forma aproximada -quizá porque no entienden que "*x* está mal" implique que hay razones suficientes para no hacer *x*, o porque entienden que las razones meramente subjetivas cuentan para que una acción sea incorrecta-, entonces quizá aceptaría un cambio de terminología. Tal vez, en lugar de "intuicionismo ético", debería llamar a la teoría que he estado defendiendo "intuicionismo evaluativo" o "intuicionismo sobre razones prácticas".

Notas

1. Mackie (1977, pp. 38-42). He cambiado el nombre del argumento de "el argumento de la rareza" a "el argumento de la rareza". La apelación de Jackson a la "plausibilidad metafísica" (2003, p. 562) es una versión de este argumento.

2. Otras cosas comúnmente rechazadas por motivos de rareza son: las mentes (si no

son idénticas a los cerebros), los universales, los objetos abstractos en general, el conocimiento a priori y el libre albedrío libertario.

3. En particular, Mackie encuentra "rara" la conexión entre las propiedades morales y las no morales, el poder motivador de las propiedades morales y nuestro conocimiento de las propiedades morales. Los discuto en la sección 8.2, en el capítulo 7 y en la sección 5.4, respectivamente.

4. Mackie (1977, p. 38).

5. Mackie (1977, pp. 31-5).

6. El esfuerzo por definir "mesa" es un ejercicio de clase; nadie lo ha conseguido. Consideremos una definición inicialmente plausible: "Una mesa es un mueble que consta de una superficie plana y nivelada y de soportes, destinados a apoyar otros objetos más pequeños" (Rand 1990, p. 41). Pero una mesa no tiene por qué contener soportes: Una vez vi una mesa en Fallingwater pegada directamente a la pared. La mesa estaba apoyada, pero ningún soporte formaba parte de la mesa. Supongamos que modificamos la definición para decir simplemente que la superficie debe estar apoyada en el suelo. La definición sigue implicando que una cama o un taburete son una mesa. Supongamos que insertamos "que no sean personas" después de "objetos". Entonces la definición implica falsamente que una mesa de operaciones no es una mesa. También implica (todavía) que los escritorios y las estanterías son mesas. Obviamente, no puedo considerar todas las modificaciones posibles, pero esto es suficiente para evidenciar que nuestra comprensión del concepto no garantiza nuestra capacidad para definirlo. De nuevo, nada de esto me hace dudar lo más mínimo de que las mesas existen.

7. La superveniencia (más concretamente, la "superveniencia fuerte") suele definirse como sigue: Las propiedades A supervienen a las propiedades *B* si: no hay dos objetos posibles que puedan diferir en sus propiedades *A* sin diferir en sus propiedades *B* (Kim 1993, p. 65). Sin embargo, la relación a la que pretendo referirme en el texto es más fuerte que esto; pretendo que la superveniencia implique también que las propiedades *A* dependen de las propiedades *B*. Obsérvese también que podríamos debilitar la tesis de la superveniencia del ámbito evaluativo a la afirmación de superveniencie global: que el mundo en su conjunto no podría haber diferido en la distribución de las propiedades evaluativas sin diferir también en la distribución de las propiedades naturales (Kim 1993, p. 68). Para simplificar, en el texto sólo considero la superveniencia fuerte.

8. Mackie (1977, p. 41).

9. Ross ([1930]1988, pp. 19-20).

10. Strawson (1949, pp. 29-30).

11. Gay (1985, p. 255) y Dancy (1993, p. 99) hacen observaciones similares.

12. Dancy (1993, pp. 100-4) discute y critica este tipo de analogía, pero hay que tener en cuenta que la analogía no tiene, como supone Dancy, nada que ver con la teoría de la propensión de la probabilidad (la física Newtoniana no hace uso de las probabilidades).

13. La mayoría de los filósofos parecen estar de acuerdo con esto, pero véase Stuart Rachels (2004) para interesantes argumentos en contra.

14. Dancy (1993, pp. 60-1,102); McNaughton (1988, pp. 200-1).

15. Es importante destacar que no pretendo refutar el particularismo del tipo defendido por Dancy o McNaughton. La noción "liberalizada" de derecho prima facie que acabo de sugerir en el texto es compatible con el particularismo, aunque no lo sigue.

16. Por ejemplo, tiene sentido intentar controlar el centro del tablero, ya que desde allí se pueden atacar más otros lugares.

17. Una matización: la propiedad de las reglas depende de la capacidad de cálculo de los jugadores. Un jugador que pudiera ver todo el árbol de partidas posibles y pudiera calcular a una velocidad ilimitada podría jugar una partida perfecta sin recurrir a estas reglas. No obstante, mantengo que las reglas son necesariamente correctas en algún sentido: es necesario que se aproximen a cómo jugaría el jugador perfecto. Tal vez ocurra algo parecido con un agente que tenga un conocimiento moral perfecto: tal vez exista una teoría moral completa y precisa que obvie la necesidad de principios de mera corrección prima facie. Pero si la hay, no la conocemos; de ahí nuestra necesidad de tales principios.

Stuart Rachels ha señalado (comunicación personal) que lo que constituye un buen juego depende de las capacidades cognitivas del oponente, en el sentido de que, siendo iguales las demás cosas, es mejor hacer una jugada a la que sea más difícil que el oponente particular vea una respuesta apropiada. Observo, como apelación a la autoridad, que Rachels (campeón de ajedrez de Estados Unidos en 1989) también ha respaldado mis observaciones sobre el ajedrez en el texto.

18. Más precisamente, *N* es la propiedad natural máxima de *A* -la propiedad formada por la unión de todas las propiedades naturales de *A*, incluyendo las propiedades negativas (como "no ser rojo"), de tal manera que nada con *N* puede tener propiedades naturales no implicadas por *N*.

19. Esto es un compuesto de Blackburn (1993, pp. 114-19) y (p. 13 7).

20. Blackburn (1993, p. 13 7); compárese con la p. 122, argumentando que cuando adoptamos una actitud hacia algo debido a ciertas propiedades, adoptaremos esa actitud hacia cualquier cosa que tenga esas propiedades.

21. Estrictamente, en lugar de "conjunto total de rasgos naturales", debería decir "rasgo natural máximo"; véase la nota 18. Si el rasgo natural máximo de *A* es la bondad, entonces *A* es bueno simpliciter, ya que entonces no puede haber otros rasgos que anulen la bondad prima facie de *A*.

22. Hay dos problemas técnicos con esta formulación del argumento: uno, supone que, o bien existen predicados infinitamente largos, o bien sólo hay un número finito de maneras de que una cosa sea buena; dos, supone en cualquier caso que las posibles maneras de que una cosa sea buena son contables. Pero debemos renunciar a estas objeciones para llegar a respuestas más interesantes al argumento de Jackson.

23. Jackson (2003, pp. 562-3).

24. Jackson (2003, p. 565); véase también (p. 562).

25. No existe un sistema formal consistente y completo capaz de derivar todas las verdades de la aritmética. Esto implica que no hay ningún algoritmo capaz de generar todas las verdades aritméticas (sin generar también contradicciones). Véase Nagel y Newman (2001) para la exposición del Teorema de Godel.

26. lowe este punto a Richard Fumerton (comunicación personal).

27. Brandt (1996, p. 4) y Abraham (1933) plantean este problema. El argumento de este párrafo es de Abraham (p. 39).

28. Tara Smith (2000, pp. 27-8).

29. Hare (1989, p. 83) acusa a Rawls de este pecado, por la confianza de Rawls en el equilibrio reflexivo.

30. Nelson (1991) hace un punto similar.

31. Esta objeción me fue sugerida por Brian Doherty y Bryan Caplan (comunicaciones personales).

32. Los científicos han determinado recientemente que el siguiente es el chiste más divertido del mundo:

> Una pareja de cazadores de Nueva Jersey está en el bosque cuando uno de ellos cae al suelo. Parece que no respira y tiene los ojos en blanco. El otro saca su teléfono móvil y llama a los servicios de emergencia. Le dice al operador: "¡Mi amigo está muerto! ¿Qué puedo hacer? La operadora, con una voz calmada y tranquilizadora, le dice: "Tranquilízate. Yo puedo ayudarle. Primero, asegurémonos de que está muerto". Se hace un silencio y luego se oye un disparo. La voz del tipo vuelve a la línea. Dice: "Vale, ¿y ahora qué?" (LaughLab s.f.).

Sin embargo, mi chiste en el texto me parece más divertido.

33. Para más información sobre este tipo de explicaciones, véase Dawkins (1989, capítulo 12), Wright (1995, especialmente el capítulo 10) y Ruse (1998, págs. 218-22).

34. Estos argumentos aparecen en Ruse (1985, pp. 237-8), Joyce (2001, capítulo 6). y quizás Wilson (2000, p. 562), aunque Wilson es oscuro. Véase también Ruse (1998, pp. 252-4).

35. Compárese la analogía de Pinker (2002, pp. 192-3) entre las matemáticas y la ética. Al igual que nuestra capacidad evolucionada para entender los números nos lleva a aprehender correctamente las verdades aritméticas independientemente existentes, quizás nuestra capacidad evolucionada para entender el bien y el mal nos lleva a aprehender correctamente las verdades morales independientemente existentes.

36. Nozick (1981, p. 345) hace esta observación. Compárese el ejemplo de James Rachels (1990, págs. 78-9) de la "mathobiology", un campo hipotético que estudiaría la base biológica del pensamiento matemático; este campo no añadiría nada a las matemáticas. Ruse (1985, pp. 236-7) parece ser consciente de este punto, pero extrañamente no le preocupa.

37. Véase más arriba, p. 103.

38. Véase Ruse (1998, p. 222).

39. Allen et al. (1977, pp. 145-6).

40. Ruse (1988, capítulo 6) discute (sin respaldar) este punto de vista.

41. Algo parecido a esta explicación se deriva de David Friedman (comunicación informal). Friedman (1994) da una explicación similar sobre los beneficios de respetar los derechos.

42. Haidt (2001) y Haidt y Hersh (2001) presentan pruebas de lo que yo llamaría

irracionalidad en el juicio moral, aunque no lo describen así. Las pruebas de Haidt y Hersh indican que los juicios morales suelen estar más causados por las emociones que por la reflexión y el razonamiento racionales. Como ejemplo, cerca de la mitad de los estadounidenses creen que las relaciones homosexuales son moralmente incorrectas (Haidt y Hersh 2001, p. 192; Gallup n.d.).

43. En 2005, el 92% de los estadounidenses pensaba que la poligamia es mala (Gallup n.d.).

44. Haidt (2001) señala que los individuos que se enfrentan a una historia de incesto inofensivo entre hermanos suelen inventar racionalizaciones para juzgar la acción como incorrecta (por ejemplo, apelando a los peligros de la endogamia). Cuando se les confronta con la falsedad de sus explicaciones (por ejemplo, recordándoles que la historia especificaba que el hermano y la hermana usaban dos formas de control de la natalidad), los sujetos suelen insistir en que la acción sigue siendo mala aunque no puedan explicar por qué.

45. Ruse (1998, pp. 145-7).

46. Véase, por ejemplo, la contribución de Richard Posner en Singer y Posner (2001): "No me siento obligado a defender esta reacción; es una intuición moral más profunda que cualquier razón que pueda darse para ella e impermeable a cualquier razón que usted o cualquiera pueda dar en contra. La pertenencia a la especie humana no es un "hecho moralmente irrelevante", como ha llegado a parecer la raza y el sexo de los seres humanos".

47. Mackie (1977, pp. 36-7).

48. Más precisamente, está determinada por las disposiciones de los hablantes competentes para aplicar la palabra -incluso aplicándola sólo en el pensamiento- a varias cosas posibles.

49. El último ejemplo es de Haidt, Koller y Dias (1993), que han estudiado las actitudes hacia acciones inofensivas y ofensivas como limpiar el retrete con la bandera, tener relaciones sexuales con un cadáver de pollo y comerse al perro de la familia cuando se muere. Haidt y Hersh (2001) han estudiado las actitudes hacia casos de incesto inofensivo. En general, muchos juicios de maldad moral parecen basarse en las emociones, son difíciles de justificar y no están relacionados con ningún daño hecho o que se crea que se va a hacer.

9

Conclusión

En los últimos ocho capítulos, he defendido el intuicionismo ético refutando las teorías alternativas y respondiendo a las principales objeciones al intuicionismo. En esencia, he argumentado:

1. Hay cinco puntos de vista amplios sobre la naturaleza de la moralidad: no cognitivismo, subjetivismo, naturalismo, nihilismo e intuicionismo.
2. El no cognitivismo, el subjetivismo, el naturalismo y el nihilismo se enfrentan a graves dificultades.
3. El intuicionismo no se enfrenta a ningún problema serio.
4. Por lo tanto, el intuicionismo es probablemente el punto de vista correcto.

Pero los argumentos que establecen los puntos (2) y (3) han sido largos y complejos. En este capítulo, hago un breve repaso de los argumentos más importantes de estas tesis. A continuación, analizo por qué el intuiciónismo ha sido impopular, y por qué la cuestión es importante. Concluyo con unos comentarios sobre por qué me considero

intuicionista.

9.1 Los fracasos de las teorías alternativas de la metaética

Las cuatro teorías alternativas de la metaética son puntos de vista opuestos sobre la naturaleza de los valores y los juicios de valor. Para facilitar la exposición, daré ejemplos de lo que dirían sobre ciertos enunciados evaluativos; en cada caso, debe entenderse que la teoría dice cosas análogas sobre todos los demás enunciados evaluativos.

9.1.1 El no cognitivismo

Los no cognitivistas éticos dicen que "*x* está mal" no afirma nada que pueda ser verdadero o falso. En su lugar, sólo expresa una emoción o emite un imperativo. Por lo tanto, es comparable a "¡Boo *x*!" o "¡No hagas *x*!".

No hay ninguna prueba lingüística que apoye esto. Todas las pruebas objetivas que se han propuesto para distinguir las afirmaciones de los imperativos o las expresiones de emoción dan como resultado que "*x* está mal" es una afirmación. Por lo tanto, considere las frases (y pseudofrases) de la siguiente tabla:

Oración normal	*Interpretaciones no cognitivistas*	
El aborto está mal.	Buu al aborto.	No abortar.
Es falso que el aborto esté mal.	Es falso que Buu al aborto.	Es falso que no aborten.
El aborto puede estar mal.	Tal vez Buu al aborto.	Tal vez no aborten.

Jon cree que el aborto está mal	Jon cree en Buu sobre el aborto	Jon cree que no debe abortar
Espero que el aborto no esté mal.	Espero no abuchear al aborto.	Espero no tener un aborto.
¿Está mal el aborto?	Buu al aborto?	¿No tienes un aborto?
Si el aborto está mal, entonces no lo hagas.	Si abuchean el aborto, entonces no lo hagan.	Si no tienes un aborto, entonces no lo hagas.
La maldad del aborto me molesta.	La locura del aborto me molesta.	Me molesta lo de no tener que abortar.

Todas las frases de la primera columna tienen perfecto sentido y son un buen inglés. La segunda y la tercera columna ofrecen intentos de interpretación no cognitivista de las mismas. Las únicas que parecen funcionar son las de la primera fila. Para cualquier frase más compleja que incluya "wrong", la interpretación no cognitivista degenera en un sinsentido. Las expresiones de la segunda y tercera columnas no sólo están mal formadas gramaticalmente, sino que no sugieren ningún significado inteligible. No tengo ni idea, por ejemplo, de lo que significaría llamar "falsa" a una emoción, ni de lo que significaría creer en un imperativo. Además, para cualquier caso no controvertido de una oración no cognitiva -como "¡Hurra por los Broncos!" o "Por favor, pásame la sal"- no surgen tales problemas; no las llamamos falsas, ni decimos que las creemos, etc.

Tampoco hay muchas pruebas psicológicas que apoyen el no-cognitivismo. El no-cognitivista puede señalar el hecho de que las emociones a menudo acompañan e incluso causan nuestros juicios morales. Pero esto es igualmente cierto en el caso de muchos juicios fácticos no morales, como las creencias de la gente sobre su propia

inteligencia, sus creencias sobre quién "ganó realmente" las elecciones presidenciales de 2000 en Estados Unidos, etc. Además, la fuerza de nuestras emociones guarda poca relación con nuestras creencias morales. Puedes saber que la ejecución de Agripina por parte de Nerón estuvo mucho más equivocada que el hecho de que tu compañero se comiera tu helado, pero aun así sentir más indignación por lo segundo que por lo primero. De hecho, es posible que no sientas nada por el destino de Agripina.

Introspectivamente, somos conscientes de hacer juicios morales, que tienen las características que normalmente tienen los juicios. Uno puede distinguir, por ejemplo, el grado de confianza de su creencia del grado de maldad que cree que tiene una acción. Uno podría creer tentativamente que el aborto es muy malo, o creer firmemente que el aborto es ligeramente malo, y estos serían estados mentales diferentes. Pero si un "juicio" moral fuera sólo una emoción o un deseo, no podría existir tal distinción; sólo existiría la fuerza de la emoción o el deseo que uno siente sobre el aborto.

Tratamos las cosas que llamamos "juicios morales" exactamente como tratamos los juicios, y no como tratamos las emociones o los deseos típicos. Podemos tener cuidado al hacer juicios morales, preguntarnos si tenemos razón, argumentar y sopesar las pruebas sobre ellos, investigar, etc. Podemos considerar argumentos dirigidos a determinar si nuestros juicios están sesgados y si son coherentes con otros juicios plausibles. No hacemos nada de esto por las emociones o deseos típicos.

Todas las pruebas razonablemente directas y objetivas están en contra del no cognitivismo: las afirmaciones y los juicios morales actúan de todas las maneras discernibles como las afirmaciones y las creencias. Lo único que el no-cognitivista tiene a su favor es la supuesta

inverosimilitud de otros puntos de vista.

9.1.2 Subjetivismo

Los subjetivistas éticos dicen que "*x* es correcto" es una afirmación sobre la actitud de alguien hacia *x*. Aquí hay tres versiones comunes del subjetivismo:

Subjetivismo individual: *x* es correcto = yo (el hablante) apruebo *x*.

Relativismo cultural: *x* es correcto = la sociedad aprueba *x*.

Teoría del mandato divino: *x* es correcto = Dios aprueba *x*.

Estas teorías se enfrentan a una serie de problemas. En primer lugar, cada una de ellas implica que si el individuo o grupo especificado cambiara sus actitudes, entonces las cosas aparentemente horribles se convertirían de repente en morales. Así, cada teoría implica una afirmación de la siguiente forma:

> Si (yo/sociedad/Dios) aprobara la tortura y el asesinato de niños por diversión, entonces sería correcto torturar y matar niños por diversión.

Pero esto no es correcto.

En segundo lugar, el subjetivismo individual implica, absurdamente, que yo soy moralmente infalible: que en general, si apruebo algo, entonces, automáticamente, es moralmente correcto. El relativismo cultural implica igualmente que la sociedad es moralmente infalible.

En tercer lugar, estas teorías hacen que el desacuerdo moral sea difícil de entender. Desde el punto de vista subjetivista individual, *A* podría estar en desacuerdo con la afirmación de *B*, "x está mal", sólo si *A* dudara de que *B* estuviera informando con veracidad de sus actitudes.

Desde el punto de vista del relativismo cultural, el desacuerdo sólo podría producirse si las partes estuvieran en desacuerdo sobre cuáles son las actitudes de la mayoría de la sociedad, y sería imposible que alguien criticara las normas sociales imperantes. Según la teoría del mandato divino, los desacuerdos morales tendrían que ser desacuerdos sobre los mandatos de Dios, por lo que los ateos y agnósticos no podrían tener opiniones morales.[1]

En cuarto lugar, todas estas teorías implican que una actitud o decisión meramente arbitraria puede hacer que una acción sea obligatoria. Todas ellas implican que no hay hechos evaluativos independientes de las actitudes del individuo o grupo especificado. Por lo tanto, el individuo o grupo especificado no puede aprobar las cosas que aprueba porque son correctas, buenas o tienen cualquier otra característica positiva. Así, estas actitudes de aprobación tendrían que ser arbitrarias. Pero, ¿cómo es posible que el hecho de que alguien apruebe arbitrariamente *x* haga que sea correcto hacerlo?

En quinto lugar, la teoría del mandato divino se enfrenta a los problemas adicionales de cómo podemos saber que Dios existe, o lo que quiere si lo hace.

Además, no hay buenos argumentos para el subjetivismo. Los principales argumentos parecen consistir en señalar que la gente tiene muchas creencias morales contradictorias o que muchas personas se forman creencias morales de forma irracional. No está claro cómo cualquiera de estas cosas podría apoyar el subjetivismo. Los argumentos tienen tanto sentido como argumentar que si tres jugadores creen, por razones emocionales, que diferentes caballos van a ganar una carrera, entonces la carrera no tendrá ningún resultado en la realidad, o tendrá diferentes resultados "relativos" a diferentes personas.

9.1.3 Naturalismo

Los naturalistas éticos sostienen dos creencias:

i) Las propiedades morales son reducibles: por ejemplo, es posible explicar qué es lo malo sin utilizar ninguna expresión evaluativa.

Hay dos variantes de esta idea:

a) Los reduccionistas *analíticos* dicen que una expresión que no contiene términos evaluativos es sinónimo de "mal".

b) Los reduccionistas *sintéticos* dicen que alguna expresión que no contiene términos evaluativos explica lo que es estar equivocado, aunque no tenga el mismo significado que "equivocado".

ii) Las verdades morales pueden conocerse, en última instancia, a partir de la observación.

El reduccionismo analítico fue refutado por G. E. Moore. Si dos expresiones tienen el mismo significado, entonces debería ser posible sustituir una expresión por la otra en cualquier oración, sin cambiar el significado de la misma. Moore observó que, para cualquier expresión no valorativa, era posible formular un enunciado o una pregunta con sentido que contuviera la palabra "bueno" en la que, intuitivamente, no se pudiera sustituir la expresión no valorativa por "bueno". Así, consideremos las tres teorías reduccionistas que se mencionan a continuación:

Definición Reduccionista	*Pregunta de Moore*	*Interpretación Reduccionista*
x es bueno = *x* promueve el placer	Es bueno el placer?	¿El placer promueve el placer?
x tiene razón =	¿Es correcto obedecer	¿Obedecer las costumbres

concuerda *x* con las costumbres sociales.	obedecer las costumbres sociales?	sociales está de acuerdo con las costumbres sociales?
x es bueno = *x* promueve la vida	Es buena la vida?	¿La vida se promueve a sí misma?

En la primera columna hay tres ejemplos de tesis reduccionistas analíticas. En la segunda columna están las preguntas que tiene sentido formular, en el discurso moral ordinario. El punto de Moore es que las preguntas de la tercera columna obviamente no significan lo mismo que las de la segunda columna. Pero deberían significar lo mismo, si las teorías reduccionistas fueran correctas. Por lo tanto, las teorías reduccionistas no son las cuentas correctas de los significados de "bueno" o "correcto".

Los reduccionistas sintéticos admiten este punto. Dicen que las propiedades morales son reducibles de forma análoga a como el agua, por ejemplo, es reducible: el agua es H_2O; sin embargo, la palabra "agua" no es sinónimo de la palabra "H_2O". Esto puede demostrarse con un argumento mooreano: la pregunta "¿Cuál es la fórmula química del agua?" no significa lo mismo que "¿Cuál es la fórmula química del H_2O?"; por tanto, "agua" no significa lo mismo que "H_2O". Sin embargo, H_2O es lo que es el agua.

Esta forma de reduccionismo se enfrenta a dos problemas principales. El primero es que las propiedades evaluativas parecen, a primera vista, radicalmente diferentes de las propiedades naturales. Ser bueno, por ejemplo, es obviamente un tipo de cosa diferente a ser redondo, pesar 200 libras o estar cargado positivamente. Esto se puede ver sobre la base de la comprensión de los conceptos de las propiedades relevantes, al igual que se puede ver, únicamente sobre la base de la comprensión del

concepto de una sinfonía y el concepto de un planeta, que el planeta Neptuno no es la Novena Sinfonía de Beethoven.

El segundo problema es que los reduccionistas sintéticos no pueden explicar el conocimiento moral. No conocemos la maldad del mismo modo que conocemos el agua: no podemos observar con los cinco sentidos que un acto es malo. Tampoco, según los reduccionistas sintéticos, podemos saber que un acto es incorrecto por intuición. Tampoco se pueden deducir proposiciones evaluativas a partir de premisas no evaluativas.

Los naturalistas contemporáneos argumentan que podemos conocer las verdades morales porque algunas verdades morales proporcionan las mejores explicaciones para ciertos hechos no morales de los que tenemos un conocimiento independiente. Por ejemplo, la maldad de Adolf Hitler supuestamente proporciona la mejor explicación de por qué ordenó la Solución Final. Pero para que esto funcione como una explicación de nuestro conocimiento moral, el naturalista tendría que mostrar que la hipótesis de que Hitler era malvado nos permite explicar hechos que no podrían explicarse de otra manera, y hacerlo sin depender de ninguna suposición *ad hoc* sobre los poderes causales del mal o similares. Así, habría que argumentar que las acciones de Hitler no podrían explicarse igual de bien apelando simplemente a hechos psicológicos no descritos moralmente, como el hecho de que tuviera un intenso odio a los judíos, un ansia de poder, etc. Los naturalistas no han podido argumentar esto.

Otra forma de ver la inutilidad del proyecto naturalista es imaginar a un individuo perverso con valores exactamente opuestos a los suyos. Esta persona podría construir "explicaciones" para los fenómenos observados exactamente paralelas a las que usted podría construir, pero utilizando "malo" donde usted utiliza "bueno", y así sucesivamente. Así,

si usted dice que las acciones de Hitler durante la Segunda Guerra Mundial se explican por su maldad, el individuo perverso podría decir que las acciones de Hitler se explican por su bondad. Si se dice que la esclavitud fue abolida porque era injusta (y la gente generalmente tiende a oponerse a la injusticia), el individuo perverso podría decir que la esclavitud fue abolida porque era justa (y la gente generalmente tiende a oponerse a la justicia). Sin duda, acusaríamos a esta persona perversa de tener valores corruptos, pero esto no podría probarse simplemente citando la capacidad de explicar los hechos observados, ya que podría ofrecer explicaciones correspondientes a cada explicación que ofreciéramos. Sus explicaciones, por supuesto, serían contraintuitivas, pero el naturalista ético no reconoce la intuición como una fuente legítima de conocimiento. En consecuencia, el naturalista no tendría forma de demostrar que una creencia moral determinada está más justificada que su opuesto exacto.

9.1.4 El nihilismo

Los nihilistas creen que todas las afirmaciones evaluativas positivas son falsas. Creen esto porque piensan que (a) las afirmaciones evaluativas son afirmaciones sobre un supuesto reino de propiedades morales objetivas no naturales, pero (b) realmente no existen tales propiedades. El aborto, diría el nihilista, no puede ser incorrecto, porque para que sea incorrecto tendría que haber cosas como valores objetivos.

Este punto de vista tiene implicaciones como las siguientes: el asesinato no es malo; el dolor más insoportable no es peor que el mayor éxtasis; el Holocausto no fue malo; etc. Es difícil saber cómo responder a este punto de vista, porque es difícil pensar en proposiciones más absurdas que algunas de las que abrazan explícitamente sus defensores.

Para apoyar su punto de vista, los nihilistas deben presentar

argumentos poderosos. Estos argumentos tendrían que partir de premisas de las que estuviéramos fuertemente convencidos; en particular, estas premisas tendrían que ser más plausibles que la afirmación de que el asesinato está mal, que el Holocausto fue malo, etc. Es difícil ver cómo el nihilista podría hacer esto.

Los argumentos reales presentados por los nihilistas han sido decepcionantes. Dependen de la supuesta inverosimilitud de la noción de moral objetiva. Pero la noción de moralidad objetiva no es inverosímil, y los argumentos contra ella han sido refutados en capítulos anteriores.

9.2 El punto de vista intuicionista

Hasta ahora hemos establecido:

1. Los enunciados evaluativos afirman proposiciones, que pueden ser verdaderas o falsas. (El no cognitivismo es falso).
2. No siempre son falsos. (El nihilismo es falso).
3. La verdad de un enunciado evaluativo no es subjetiva ni relativa. (El subjetivismo es falso).
4. Los hechos evaluativos no pueden reducirse a hechos no evaluativos, ni pueden conocerse únicamente a partir de la observación. (El naturalismo es falso.)

De estos puntos se deduce que:

5. Algunos enunciados evaluativos son verdaderos. (De 1, 2.)
6. Algunos enunciados evaluativos son objetivamente verdaderos. (De 3, 5.)
7. Hay hechos irreductibles, objetivos, evaluativos, que no pueden ser

conocidos a partir de la observación. (De 4, 6.)[2]

(7) se acerca a una definición del intuicionismo ético. Sólo le falta una declaración positiva de cómo se conocen los hechos evaluativos, a la que ahora nos referiremos.

9.2.1 Intuición y conocimiento moral

Algunos principios básicos sobre lo bueno, lo malo, lo correcto y lo incorrecto son evidentes, como los siguientes:

> El sufrimiento es malo.
>
> Si *A* es mejor que *B*, y *B* es mejor que *C*, entonces *A* es mejor que *C*.
>
> Si una acción es mala, entonces cualquier acción cualitativamente indistinguible (en idénticas circunstancias) también es mala.
>
> Ninguna persona es culpable de una acción que no haya realizado.

Estos principios son evidentes en el mismo sentido que los siguientes:

> Nada es a la vez rojo y verde.
>
> Si *A* está junto a *B*, entonces *B* está junto a *A*.
>
> El amarillo es un color.
>
> Todo es idéntico a sí mismo
>
> "Sócrates es un hombre" y "Todos los hombres son desconsiderados" implican conjuntamente "Sócrates es desconsiderado".
>
> El tiempo es unidimensional.
>
> Siete es más que dos.

Estamos justificados en creer estas proposiciones por la misma razón

que estamos justificados en creer las observaciones del mundo externo: a saber, porque parecen ser el caso y no tenemos motivos serios para dudar de ellas. El principio del *Conservadurismo Fenoménico* dice: si a uno le parece que *p*, entonces tiene al menos una justificación *prima facie* para creer que *p*. Este principio se presupone en todos los juicios y razonamientos normales, ya que los juicios que uno hace están determinados por lo que a uno mismo le parece que es el caso. Los argumentos que una persona acepta dependen también de lo que le parece que es válido y que tiene premisas verdaderas. Incluso los escépticos que afirman que no sabemos nada lo dicen sólo porque les parece que es así (o les parece que se deduce de cosas que les parecen que son así).

Nos formamos creencias justificadas sobre el mundo físico a partir de las apariencias sensoriales (también llamadas "experiencias perceptivas"). Del mismo modo, nos formamos creencias justificadas sobre ciertas verdades abstractas, incluidos los principios autoevidentes mencionados anteriormente, a partir de apariencias *intelectuales* ("intuiciones"). Las intuiciones son estados mentales en los que algo parece ser el caso tras la reflexión intelectual (por oposición a la percepción, la memoria o la introspección), antes de la argumentación.

Cuando percibimos objetos físicos, no aprehendemos primero las experiencias perceptivas y luego inferimos que hay objetos físicos que se corresponden con ellas. Por el contrario, primero somos conscientes de los objetos físicos, directamente. Nuestras experiencias perceptivas no son más que la forma que adopta esa conciencia; constituyen (en parte) nuestra conciencia del mundo físico. Del mismo modo, cuando conocemos las verdades morales intuitivamente, no aprehendemos primero las intuiciones y luego inferimos que hay hechos morales que les corresponden. En cambio, somos directamente conscientes de los

hechos morales. Las intuiciones son, en realidad, la forma que toma esa conciencia; constituyen (en parte) nuestra conciencia de las verdades morales.

Las intuiciones morales pueden verse afectadas por prejuicios, al igual que casi todas las creencias. Sin embargo, las intuiciones morales no son creencias morales, ni están simplemente causadas por nuestras creencias morales. Esto se demuestra por el hecho de que, cuando nos enfrentamos a casos que nunca hemos considerado previamente, a menudo tenemos intuiciones morales que entran en conflicto con nuestras teorías lnorales. Así, incluso los utilitaristas experimentan la intuición de que estaría mal matar a una persona sana para distribuir sus órganos a otras cinco personas que necesitan trasplantes, aunque los utilitaristas no creen en esta intuición. En algunos casos, tenemos dificultades incluso para elaborar una teoría que dé cabida a nuestras intuiciones.

9.2.2 La motivación moral

¿Por qué debemos ser morales? Según la teoría más popular de las razones prácticas, la única buena razón, así como la única motivación, que alguien puede tener para hacer cualquier cosa es que satisfaga sus deseos. Esta es la teoría humeana de las razones.

La teoría humeana es demasiado simplista. Hay al menos dos tipos diferentes de motivación que uno puede tener: uno puede estar motivado por deseos, y uno puede estar motivado por ciertos tipos de creencias. El humeano tiene razón al señalar que la mayoría de los tipos de creencias no pueden motivar la acción por sí mismos. Por ejemplo, mi creencia de que no tengo galletas no puede moverme por sí misma a comprar más galletas en la tienda, ya que la creencia de que no tengo galletas no me dice por sí misma si tener galletas es deseable o no. Pero hay algunos

tipos de creencias que son diferentes, como las creencias prudenciales (creencias sobre lo que le conviene a uno) y las creencias morales. Estas creencias, por su propia naturaleza, nos dicen lo que es deseable y lo que no. La creencia de que me he quedado sin galletas puede ser incapaz de motivar la acción, pero la creencia de que me he quedado sin galletas y que sería bueno tener algunas galletas, sí puede hacerlo.

Estas creencias evaluativas hacen que ciertas acciones sean *prima facie* racionales. La creencia de que *x* es en el interés a largo plazo proporciona una buena razón para hacer *x*. Además, esta razón tiene prioridad sobre cualquier deseo que uno pueda tener en sentido contrario, en el sentido de que incluso si uno desea más fuertemente hacer otra cosa, seguiría siendo racional hacer *x*. Esto es porque la creencia prudencial (si se forma racionalmente) ya tiene en cuenta la evaluación de uno de la importancia relativa de sus deseos presentes; por lo tanto, los deseos presentes no pueden pesar en contra de ella. Del mismo modo, la creencia de que *x* es moralmente necesaria proporciona una buena razón para hacer *x*. Esta razón tiene prioridad no sólo sobre los deseos, sino también sobre las creencias prudenciales. Esto se debe a que la creencia moral (si se forma racionalmente) ya tiene en cuenta la evaluación de la importancia relativa de los propios deseos e intereses; por tanto, los deseos e intereses no pueden pesar en contra.

Pero aunque las creencias prudenciales y morales tienen este tipo de autoridad racional, las personas no siempre actúan racionalmente: hay casos de debilidad de la voluntad, en los que una persona actúa por algún deseo presente en lugar de por su juicio sobre lo que realmente le interesa, o sobre lo que es moralmente correcto. Tanto la debilidad de la voluntad como la fuerza de la voluntad son posibles, porque el peso de los motivos racionales para la acción no es conmensurable con la fuerza de los deseos; esto se debe a que un motivo racional es una creencia, y

no tiene sentido preguntarse si una creencia es "más fuerte" o "más débil" que un deseo. Este punto de vista deja al menos espacio para el libre albedrío, algo que la teoría humeana parece incapaz de hacer.

La teoría humeana de las razones es inferior en varios aspectos al relato que acabamos de esbozar. En primer lugar, la teoría humeana no nos ofrece ninguna comprensión del fenómeno de la debilidad de la voluntad. Desde el punto de vista humeano, cuando dos motivos entran en conflicto, ya que todos los motivos provienen del deseo, la única cuestión es: ¿qué deseo es más fuerte? El deseo más fuerte siempre se llevará a cabo. Además, dado que todos los buenos motivos para actuar provienen del deseo, será racional actuar según el deseo más fuerte. No hay nada en esta historia que se parezca a la "debilidad", que explique cómo los agentes pueden actuar irracionalmente, o que explique la conveniencia de las reacciones de remordimiento o de culpa.

En segundo lugar, los humeanos no pueden explicar qué es lo interesante o particularmente importante de la moral. Desde el punto de vista humeano, lo más que puede decirse de la moralidad es que sus reglas a menudo nos ayudan a satisfacer ciertos deseos, como nuestros deseos por el bienestar de los demás. Pero tenemos muchos otros deseos que pueden entrar en conflicto con la motivación moral, como los deseos causados por la envidia, la ira y el odio; y desde un punto de vista humeano, no hay ninguna razón por la que los primeros deseos deban ser de alguna manera privilegiados sobre los segundos. La moralidad, por lo tanto, sólo sirve a una subclase arbitrariamente seleccionada de nuestros deseos, por los que no tenemos más razón para interesarnos que por cualquier otro deseo que tengamos. El punto de vista racionalista y antihumeano evita este problema apelando a hechos evaluativos objetivos: el hecho de que las acciones provocadas por la envidia y el odio son generalmente malas, por ejemplo.

En tercer lugar, la concepción Humeana de las razones no puede dar cabida a ejemplos paradigmáticos de acciones racionales e irracionales, declarando en cambio que ninguna acción puede ser irracional. Por ejemplo, el humeano se ve obligado a decir que los siguientes son ejemplos de acciones igualmente racionales:

> *Odiseo*: Odiseo sabe que cuando se acerque a las sirenas sentirá un intenso deseo de unirse a ellas, pero ahora desea no unirse a ellas. Odiseo no da importancia a su deseo futuro y, en cambio, se hace atar al mástil para no poder actuar en el futuro según ese deseo.
>
> *El estudiante fiestero*: Un estudiante decide pasar la noche de fiesta, a sabiendas de que eso le hará suspender un importante examen al día siguiente. Aunque sabe que mañana deseará haber pasado la noche estudiando, da poca o ninguna importancia a este deseo futuro y, en cambio, actúa según su deseo inmediato de salir de fiesta.

Para un humeano, las razones para actuar que uno tiene ahora deben estar determinadas por los deseos que uno tiene ahora. La única razón por la que uno se preocuparía por sus deseos futuros sería si tuviera un deseo presente de que sus deseos futuros fueran satisfechos, pero no hay ninguna razón por la que uno deba tener tal deseo presente. El humeano tiene razón en el caso de Odiseo: todos podemos estar de acuerdo en que Odiseo es racional al actuar para satisfacer su deseo presente, en lugar de su deseo futuro de estar con las sirenas. Pero el humeano está igual de equivocado en el caso del Estudiante Fiestero: éste es un caso paradigmático de comportamiento irracional. El racionalista tiene una explicación lista: El deseo futuro de Odiseo de estar con las sirenas será un deseo irracional, porque será un deseo de lo que es claramente malo; el deseo futuro del estudiante de aprobar el examen, sin embargo, será un deseo perfectamente racional. Así, es racional que Odiseo descarte su deseo futuro, pero irracional que el estudiante haga lo mismo. Un

humeano no puede aceptar esta explicación, ya que los humeanos niegan que los deseos últimos puedan ser evaluados por la razón.

En cuarto lugar, un humeano ni siquiera puede explicar por qué deberíamos preocuparnos de que nuestros juicios morales sean coherentes. La razón habitual para querer creencias coherentes es que las creencias coherentes tienen más probabilidades de ser verdaderas. Pero para un humeano, los principios morales -en la medida en que tenemos alguna razón para prestarles atención- deben reflejar nuestros deseos y sentimientos, más que nuestras creencias. Los deseos difieren fundamentalmente de las creencias en que, mientras que pretendemos que nuestras creencias se ajusten a la realidad, no pretendemos que nuestros deseos se ajusten a la realidad, sino que pretendemos que la realidad se ajuste a nuestros deseos. Así, mientras que sería irracional seguir manteniendo una creencia que se ha demostrado que es poco probable que sea cierta, no sería en absoluto irracional tener un deseo que se ha demostrado que es poco probable que se satisfaga. Por esta razón, no se puede argumentar que debamos procurar tener deseos coherentes con el argumento de que tales deseos tienen más probabilidades de coincidir con la realidad. Si no se exige coherencia a los deseos, tampoco se exigiría coherencia a los principios morales.

A pesar de su popularidad, la teoría humeana fracasa estrepitosamente como explicación de la racionalidad práctica: clasifica erróneamente los casos paradigmáticos de acciones irracionales como racionales; reduce la "acción racional" a cualquier cosa que un agente elija realmente; y no puede explicar la debilidad de la voluntad, el libre albedrío o la importancia de la moralidad. La única justificación que suele ofrecerse para la teoría humeana es una apelación a la intuición: parece que sólo los deseos pueden motivar. Pero esta intuición depende de que consideremos sólo ciertos tipos de casos. Cuando consideramos

las creencias evaluativas, no es en absoluto intuitivo que tales creencias no puedan motivar.

9.3 Algunas objeciones al intuicionismo

Se han planteado diversas objeciones al intuicionismo ético. A continuación se presentan las cuatro más comunes.

9.3.1 "No se puede demostrar que la intuición sea fiable"

Algunos dicen que no podemos conocer las verdades morales por medio de la intuición a menos que podamos verificar primero que la intuición en general es fiable. A los que hacen esta objeción no les satisfaría el uso de las intuiciones para apoyar otras intuiciones -demostrando, en otras palabras, que uno tiene un sistema de intuiciones internamente coherente-. Tampoco se darían por satisfechos demostrando que sus intuiciones coinciden con las de otras personas. Más bien, los objetores buscan una verificación de la fiabilidad de la intuición que no dependa de la intuición de nadie en ningún momento.

No se puede presentar tal verificación. Pero esto no es un problema, al menos por dos razones. En primer lugar, porque tampoco existe tal verificación para la percepción sensorial, la memoria, la introspección o la razón en general. Nadie puede demostrar que la percepción de los sentidos es generalmente fiable, utilizando sólo pruebas no perceptivas. Nadie puede probar que la memoria es generalmente fiable, utilizando sólo la información recogida en el instante presente en el tiempo. Nadie puede demostrar que la razón es fiable, sin usar la razón. Pero no declaramos ilegítimas todas estas formas de cognición. Hacerlo nos llevaría al pozo del escepticismo universal. Parece, pues, que la presente objeción se basa en un doble rasero epistemológico: el objetor impone a

la intuición exigencias que no se plantearían a ninguna otra fuente fundamental de conocimiento.

En segundo lugar, la demanda nunca estuvo justificada en primer lugar. ¿Por qué cualquier proceso de cognición debería exigir un segundo proceso cognitivo dirigido al primero? ¿Por qué no sostener, en cambio, de acuerdo con el principio del conservadurismo fenoménico, que nuestras facultades cognitivas se presumen inocentes hasta que se demuestre su culpabilidad? Esto evita el dogmatismo al dejarnos abiertos a la posibilidad de revisar nuestras creencias iniciales si aparecen nuevas pruebas que las desmientan, pero también evita el escepticismo extremo que nunca nos permitiría ningún punto de partida para nuestros sistemas de creencias.

La exigencia de verificación de la fiabilidad parecería adecuada si concibiéramos la intuición (y la percepción, la memoria, etc.) como algo que nos proporciona una conciencia inmediata de los estados mentales, a partir de los cuales intentamos inferir conclusiones sobre el mundo real. Tal inferencia requeriría la suposición previa de una correlación entre tales estados mentales y los estados del mundo real, una suposición que requeriría justificación. Pero, como he insistido más arriba, esto sería un malentendido del papel de la intuición; deberíamos concebir las intuiciones, al igual que las experiencias perceptivas y las experiencias de memoria, como meros vehículos de la conciencia de hechos que existen independientemente de ellas.

9.3.2 "Los intuicionistas no pueden explicar el desacuerdo"

Algunos afirman que el intuicionismo hace imposible entender cómo pueden producirse el error y el desacuerdo moral. Los que insisten en esta objeción a menudo parecen tener en mente una caricatura infalible del intuicionismo que probablemente no se corresponde con las

opiniones de ningún intuicionista real. De hecho, ya es bastante difícil encontrar filósofos que reivindiquen la infalibilidad sobre cualquier cosa, y mucho menos sobre la ética.

Los que evitan caricaturizar injustamente el intuicionismo argumentan simplemente que hay más errores y desacuerdos de los que cabría esperar si el intuicionismo fuera cierto. Hay al menos cuatro razones por las que se equivocan. En primer lugar, suelen exagerar la cantidad de desacuerdos y errores que existen. Mientras que hay un desacuerdo generalizado sobre cuestiones como el aborto, la discriminación positiva y la pena capital, no hay disputas serias sobre la conveniencia de cosas como el asesinato, la violación y el robo a mano armada. El primer tipo de cuestiones recibe mucha más atención que el segundo, simplemente porque no solemos discutir lo que todo el mundo sabe. Esto lleva a una visión exagerada de la prevalencia de la controversia.

En segundo lugar, el error humano y el desacuerdo son comunes con respecto a muchas cuestiones objetivas y fácticas. La gente discrepa profundamente sobre cosas como quién disparó a JFK, si un determinado jugador de baloncesto es mejor que otro, si un individuo concreto es más inteligente que otro, si existe Dios, y prácticamente todas las cuestiones importantes de la filosofía. Además, a lo largo de la historia, casi todos los que han tenido alguna opinión al respecto han tenido opiniones radicalmente erróneas sobre cosas como: las causas de la salud y la enfermedad, la estructura del cosmos, el origen de la especie humana, la existencia de espíritus incorpóreos, la composición del mundo físico y el funcionamiento de la economía, por nombrar algunas. Pero nadie concluye que todas estas cuestiones sean "subjetivas" ni que los seres humanos carezcan de facultades cognitivas legítimas capaces de conocerlas.

En tercer lugar, en general, incluso cuando disponemos de un medio de cognición válido, los seres humanos están sujetos a una gran variedad de causas de error. Entre ellas se encuentran el sesgo, el descuido, el error de cálculo, la confusión, la información falsa o incompleta y las falacias lógicas, entre otras. Es un filósofo ingenuo el que piensa que los procesos intelectuales se desarrollan casi siempre de forma fluida y limpia, sin defectos ni influencias no racionales. No hay ninguna razón obvia para que los problemas que conducen a errores y desacuerdos sobre tantas otras cosas no se produzcan cuando la gente piensa en los valores. Aquí, al igual que con la objeción anterior, se está aplicando un doble rasero epistemológico: casi nadie considera que la validez de otros medios de cognición, como el razonamiento empírico, quede refutada por la existencia de errores y desacuerdos generalizados, ni por nuestra susceptibilidad a errores e influencias no racionales de diversa índole. Sólo la intuición se ve impugnada por estos hechos. Por supuesto, los hechos justifican que seamos cautelosos en nuestras afirmaciones éticas, al igual que debemos ser cautelosos en nuestras afirmaciones sobre los otros asuntos de controversia mencionados anteriormente.

En cuarto lugar, si consideramos en qué tipo de cuestiones la gente suele estar más en desacuerdo y cometer más errores, es fácilmente predecible que las cuestiones morales estarían entre ellas. Porque la gente tiende a estar en desacuerdo y a cometer más errores sobre (a) los asuntos sobre los que tienen un fuerte sesgo, incluyendo, pero no limitándose, a aquellos en los que están en juego sus propios intereses, (b) los asuntos en los que sus creencias se basan en gran medida en el adoctrinamiento cultural, (c) los asuntos en los que sus creencias se basan en gran medida en la religión, y (d) los asuntos filosóficos en general. Esto es cierto independientemente de que las cuestiones en cuestión sean evaluativas. Dado que las cuestiones morales entran en estas cuatro categorías, es de esperar que el desacuerdo y el error sean

comunes con respecto a las cuestiones morales. Por lo tanto, no hay nada sorprendente en el fenómeno del desacuerdo moral que exija una interpretación antirrealista.

9.3.3 "Los intuicionistas no pueden resolver los desacuerdos"

Algunos objetan que el intuicionismo no nos proporciona ningún método razonable para resolver los desacuerdos morales. La forma más plausible inicialmente de la objeción afirma que hay muchas diferencias de intuiciones, y que los intuicionistas no tienen ningún método para adjudicar entre intuiciones contrarias.

Esta objeción falla por varias razones. En primer lugar, algunas intuiciones éticas pueden coincidir mejor o peor con el resto de nuestras intuiciones. Así, los filósofos morales suelen argumentar sus posiciones sobre cuestiones morales controvertidas apelando a nuestras intuiciones sobre casos (supuestamente) análogos. También suelen apelar a intuiciones sobre qué tipo de factores pueden ser moralmente relevantes. Éstas son sólo dos de las formas en las que se puede apelar a otras intuiciones en cuestiones controvertidas.

En segundo lugar, podemos utilizar nuestro conocimiento de los factores que tienden a afectar a la fiabilidad de las creencias de las personas para determinar qué intuiciones tienen más probabilidades de ser correctas. Podemos tener razones para creer que ciertas intuiciones morales están indebidamente influidas por los prejuicios, el adoctrinamiento, la religión, etc. Esto es cierto, por ejemplo, para muchas actitudes morales relacionadas con la sexualidad y la reproducción.

En tercer lugar, incluso si algunas disputas morales resultan ser irresolubles, no está claro cómo se supone que esto refuta el

intuicionismo. Hay muchas disputas, tanto sobre asuntos morales como no morales, que parecen a todos los efectos prácticos irresolubles. ¿Por qué deberíamos suponer que, con la teoría correcta, todas las disputas morales serían resolubles? Una vez más, la objeción al intuicionismo se basa en un doble rasero: para ninguna otra fuente de conocimiento fundamental exigimos un algoritmo para resolver todas las disputas.

En cuarto lugar, la objeción se basa en otro doble rasero: ninguna otra teoría metaética se somete normalmente a la exigencia de proporcionar un método para resolver todas las disputas morales, y ninguna otra teoría metaética puede hacerlo. Ni los no cognitivistas, ni los subjetivistas, ni los naturalistas han producido un método para asegurar el acuerdo sobre la moralidad del aborto, la pena capital y similares.

Por último, todas las versiones de la objeción del "desacuerdo" parecen autorrefutarse. Los que presentan la objeción afirman que su objeción en sí misma, junto con la conclusión metaética que apoya, es objetivamente correcta. Sin embargo, muchas personas inteligentes siguen estando en desacuerdo con el argumento del desacuerdo y con cada teoría metaética. ¿Cómo pueden los antirrealistas o los naturalistas explicar este desacuerdo? ¿Y cómo pueden resolverlo? Las perspectivas de resolver este desacuerdo no parecen mejores que las de resolver el desacuerdo sobre la moralidad del aborto.

9.3.4 "El intuicionismo es raro"

Muchos consideran que el intuicionismo ético es demasiado "raro" o "misterioso" para ser aceptado. Esta crítica tiene dos partes: (a) los hechos morales irreductibles son un tipo de hecho extraño; (b) la intuición es una forma extraña de conocer.

No está claro en qué consiste esta crítica. ¿Existe una propiedad

objetiva de "rareza"? Si no hay valores objetivos, entonces es dudoso que haya rareza objetiva. Pero si no la hay, entonces es difícil ver cómo la supuesta rareza del intuicionismo contaría como evidencia en su contra. Si, por ejemplo, "*x* es raro" sólo expresa una reacción subjetiva hacia *x*, entonces no proporciona ningún argumento contra la realidad de .

Si hay rareza objetiva, ¿en qué consiste? Consideremos tres interpretaciones. En primer lugar, quizá "raro" signifique "muy diferente de la mayoría de las cosas". Los hechos morales son muy diferentes de otros tipos de hechos, y la intuición es muy diferente de otros tipos de conciencia. Pero, ¿cómo podría esto demostrar que los hechos morales o la intuición moral no existen? Vivimos en un mundo diverso: hay muchas cosas que conocemos que son muy diferentes entre sí.

En segundo lugar, quizá "raro" signifique "contraintuitivo". Pero la existencia de hechos morales no es contraintuitiva. El discurso moral ordinario, en el que todos participamos a diario, está comprometido con la existencia de hechos morales. Y como muestran mis argumentos contra el subjetivismo, y como es comúnmente aceptado incluso entre los antirrealistas, normalmente pensamos en estos hechos como objetivos. Sería difícil argumentar que es el intuicionista quien adopta una posición contraintuitiva.

En tercer lugar, quizá "raro" signifique "mal entendido". Pero, de nuevo, carecemos de cualquier razón clara para pensar que las cosas que son raras en ese sentido no existen. Tampoco hay ninguna razón obvia para pensar que las propiedades morales o la intuición moral se entienden peor que muchas otras cosas cuya existencia aceptamos, como la conciencia, los números, las propiedades disposicionales, las posibilidades, el tiempo, el espacio, los campos, las propiedades estéticas y las proposiciones.

El argumento de la rareza, por tanto, es tan oscuro como influyente. Está muy poco claro qué es la rareza, o cómo la acusación de rareza proporciona una objeción a una teoría. El intuicionismo bien puede ser una posición "rara" en algún sentido, pero si lo es, esto no parece proporcionar ninguna prueba contra su verdad.

9.4 La revuelta contra los valores: cómo el intuicionismo perdió el favor

Con esto concluye mi argumento oficial y filosófico a favor de la intuición ética. Ahora voy a permitirme hacer especulaciones psicológicas y sociológicas sobre la resistencia al intuicionismo ético. Esta especulación no proporcionará pruebas a favor del intuicionismo; sin embargo, puede tener cierta relevancia lógica, ya que sugerirá que la mayoría de los que han rechazado el intuicionismo han formado sus creencias de una manera poco fiable.

9.4.1 ¿Fue el intuicionismo refutado racionalmente?

El intuicionismo ético era popular entre los filósofos de habla inglesa a principios del siglo XX. A mediados de siglo, había sufrido una dramática pérdida de apoyo en su favor, que continúa hasta nuestros días. ¿Por qué? Una explicación, que se ajusta a los antirrealistas, es la siguiente: "El intuicionismo fue la primera visión de la ética que se le ocurrió a la gente, en los albores de la filosofía analítica, porque es la visión más simple e ingenua. El intuicionismo es similar a la visión de la ética de Platón, que, no por casualidad, surgió en los albores de la filosofía occidental. A medida que la gente reflexionaba más, se dio cuenta rápidamente de los diversos defectos fatales del sistema y lo rechazó racionalmente en favor de teorías más sofisticadas y

defendibles".

Es de suponer que así es como muchos filósofos contemporáneos ven esa parte de la historia intelectual. Pero yo creo que eso es una absoluta fantasía. En mi opinión, los movimientos generales de la historia de la filosofía se parecen mucho más a cambios de temperamento y modas intelectuales que a movimientos racionales que se alejan de teorías inviables y se acercan a otras más objetivamente defendibles. No puedo apoyar este juicio de manera general aquí, ya que requeriría una discusión de varias cuestiones filosóficas importantes y episodios en la historia de la filosofía. Pero sí puedo apoyar esta afirmación en lo que respecta al intuicionismo ético y su caída de apoyo filosófico. El intuicionismo no fue rechazado porque se considerara inviable. Primero se rechazó, y luego se racionalizó el rechazo con un puñado de objeciones de poca monta, incluyendo las que se discutieron en la sección anterior. Que esto fue lo que realmente ocurrió lo demuestran dos hechos destacados.

En primer lugar, está el hecho de que la mayor parte de las objeciones prominentes al intuicionismo se basan en un doble rasero epistemológico, planteando exigencias a la intuición y al intuicionismo que no se impondrían a ningún otro medio básico de cognición, ni a ninguna otra teoría metaética. Este tipo de razonamiento es característico de las racionalizaciones. Y una de las principales objeciones parece descansar en nada más que un sentido subjetivo de "rareza", que es difícil de distinguir de un mero prejuicio.

En segundo lugar, está el hecho de que la mayoría de las objeciones al intuicionismo, si no todas, están sujetas a respuestas bastante simples y naturales, que no requieren modificaciones significativas de la teoría; sin embargo, los críticos del intuicionismo no han dedicado prácticamente ninguna energía a considerar tales respuestas. En cambio,

las teorías metaéticas más de moda han sido objeto de numerosas y elaboradas modificaciones y variaciones técnicas a lo largo de décadas, diseñadas para mostrar cómo pueden escapar de las objeciones aparentemente devastadoras que se encuentran muy cerca de la superficie. En resumen, los metaéticos de los últimos cincuenta años nunca han prestado al intuicionismo ético la misma atención que a otras teorías.

Por supuesto, niego que el intuicionismo haya sido refutado, ya que creo que la teoría es verdadera y que la he demostrado en las páginas anteriores de este libro. Pero incluso si piensas que estoy equivocado en esto, deberías estar de acuerdo en que he presentado una defensa para el intuicionismo que es tan razonable, al menos, como la defensa que se ha presentado para algunas de las otras teorías en metaética. Pero en la literatura ética de los últimos cincuenta años, el intuicionismo se ha tomado mucho menos en serio y se le ha prestado mucha menos atención que a cualquiera de las otras cuatro teorías metaéticas. Así que, sea o no cierto el intuicionismo, parece claro que la oposición al mismo ha sido desproporcionada con respecto a cualquier análisis objetivo del peso de los argumentos. Se requiere alguna otra explicación para la caída en desgracia del intuicionismo.

Creo que el intuicionismo fue rechazado porque no encaja con el espíritu de la era moderna. Hay al menos tres tendencias importantes de la cultura occidental moderna que son relevantes aquí: el cinismo, la corrección política y el cientificismo.

9.4.2 Cinismo

En el siglo XX, todo lo que antes se tenía en alta estima -la naturaleza humana, la ciencia, la verdad, el conocimiento objetivo, el arte, la filosofía, la religión, la moral- tuvo que ser desacreditado. Los

psicólogos trataron de demostrar que el hombre, antes considerado "el animal racional", era fundamentalmente irracional. Los freudianos nos encontraron llenos de lujuria e impulsos socialmente destructivos; los marxistas encontraron que la historia humana estaba dominada por la codicia y la explotación; los sociobiólogos encontraron que todos los aspectos de la vida humana estaban dominados por impulsos "egoístas" de reproducción. Los tres grupos consideraron que los seres humanos se autoengañaban masivamente y que búsquedas supuestamente espirituales como el arte, la religión y la filosofía eran tapaderas de algo más superficial y mucho menos noble de lo que parece. Muchos han comentado lo diminuta e insignificante que es la humanidad, lo tontos que fuimos al creernos el centro del universo y lo ignorantes e indefensos que somos. Muchos filósofos, que antes veían su disciplina como una búsqueda racional de las verdades fundamentales sobre la naturaleza de la realidad y nuestro lugar en ella, ahora niegan que la filosofía pueda decirnos algo de esas verdades, si es que existen. Una influyente escuela de pensamiento declaró que los filósofos no podían hacer más que discutir cómo la gente utiliza las palabras, y que la metafísica, la teología y la ética carecían de sentido. Otros filósofos declararon que no existe la verdad, que nunca podremos conocerla, que somos fundamentalmente irracionales y/o que la objetividad es imposible. Los intelectuales estadounidenses atacaron a los que en su día fueron los antepasados más venerados de su país -como Cristóbal Colón, Thomas Jefferson y George Washington- como hipócritas, fanáticos y explotadores; también atacaron la Declaración de Independencia y la Constitución. Y los intelectuales occidentales atacaron al cristianismo más que a ninguna otra religión. Los pensadores modernos han tratado de derribar los héroes de su propia sociedad, su religión, su filosofía, su sistema económico y político, en definitiva, sus valores.

No pretendo explicar todos los desarrollos intelectuales mencionados

en el párrafo anterior; confío en la familiaridad del lector con la cultura intelectual contemporánea para reconocer la mayor parte de lo que acabo de mencionar. Tampoco pretendo discutir sus méritos intelectuales; en algunos casos, las actitudes pueden estar justificadas, mientras que en otros no lo están. Lo que quiero decir es que, independientemente de lo que se diga de ellos, todos esos desarrollos culturales tienen al menos una cosa en común: el cinismo. La cultura occidental de los últimos cincuenta años debe ser una de las más cínicas de la historia del mundo. Y ese cinismo a ultranza ha hecho imposible la aceptación generalizada del intuicionismo ético en nuestra cultura, sean cuales sean los méritos intelectuales de la teoría. Una cultura que insiste en que los seres humanos son fundamentalmente egoístas e innobles nunca aceptará mi explicación de la motivación moral. Una cultura que declara ilusorias las nociones de verdad, conocimiento y objetividad nunca aceptará mi versión del conocimiento moral. Y una cultura que busca derribar todos los valores nunca aceptará el realismo moral. El subjetivismo, el no cognitivismo y el nihilismo han sido populares, en definitiva, porque ofrecen a la mente moderna el perverso placer de desacreditar la moral.

9.4.3 La corrección política

La corrección política tiene dos aspectos estrechamente relacionados: Primero, el imperativo de evitar ofender, incluso a las personas más sensibles. Así, llamamos a los basureros "ingenieros sanitarios". A los estúpidos los llamamos "retrasados mentales". ("Retrasado" era originalmente un eufemismo para "estúpido", pero desde entonces ha llegado a parecer aún más despectivo que "estúpido"; de ahí la necesidad de un nuevo eufemismo). Lo que antes era el vicio de la embriaguez se ha convertido en la "enfermedad" del alcoholismo. Es sobre todo a los grupos, más que a los individuos, a los que hay que tener cuidado de ofender.

En segundo lugar, la corrección política implica un igualitarismo extremo. Para ser políticamente correcto, hay que evitar decir que alguien es mejor que otro en cualquier aspecto significativo. Esto es un corolario de la primera regla: negar la igualdad podría ofender a los que son juzgados como peores.

Los realistas morales parecen violar estas dos restricciones. Si existen requisitos morales reales, es de suponer que no todo el mundo los satisface: a veces, la gente hace el mal. Algunas personas podrían incluso ser juzgadas como malvadas, si perpetran males suficientemente graves. En cualquier caso, si hay virtudes y vicios, entonces algunas personas resultarán tener más virtudes y menos viciosas que otras: ¿no tendremos entonces que llamar a estas personas "moralmente mejores" que las demás? Pero decir tal cosa seguramente ofendería a los que se consideran peores. De hecho, hay pocas formas más fiables de ofender a una persona que llamarla inmoral.

Además, si hay hechos objetivos de cualquier tipo, entonces algunas creencias -las que corresponden a los hechos- son presumiblemente mejores que otras. Pero, de nuevo, decir esto seguramente ofendería a quienes sostienen las creencias "inferiores" (es decir, falsas). Teniendo en cuenta este punto y el anterior, podría resultar que algunas culturas son moralmente peores que otras, ya que podrían tener creencias morales falsas y llevar a cabo prácticas inmorales. Si juzgamos, por ejemplo, que los hombres y las mujeres tienen realmente los mismos derechos, ¿no nos veríamos obligados a concluir que la cultura islámica es mala en un aspecto muy significativo: adopta creencias morales falsas y trata a las mujeres de forma incorrecta? Decir algo así seguramente ofendería a los musulmanes. ¿Pero cómo podríamos evitar decir tales cosas, si creyéramos en valores objetivos? Parece que, para ser políticamente correctos, debemos ser antirrealistas.

Mi propósito aquí no es discutir la cultura musulmana, la naturaleza del mal, el alcoholismo, etc. Mi propósito es explicar el ascenso del antirrealismo moral en los tiempos modernos: la popularidad del antirrealismo, creo, se debe en gran parte a la percepción de que el realismo moral nos obligaría a adoptar posiciones incómodas que causarían ofensa a ciertos grupos.

9.4.4 El cientificismo

Para explicar qué es el cientificismo, no puedo hacer nada mejor que citar a Peter van Inwagen:

> *El cientificismo, tal y como yo lo llamo, es una especie de respeto exagerado por la ciencia, es decir, por las ciencias físicas y biológicas, y el correspondiente menosprecio de todos los demás ámbitos del quehacer intelectual. No es necesario señalar que el cientificismo es la ideología primaria de nuestra época. No es necesario señalar que las ilusiones que engendra el cientificismo están tan extendidas y son tan insidiosas que es prácticamente imposible conseguir que alguien que esté sometido a ellas considere la posibilidad de que puedan ser ilusiones.*[3]

La ciencia moderna es algo excelente: nos ha proporcionado una comprensión sin precedentes del mundo y una capacidad sin precedentes para modificar el mundo para satisfacer nuestras necesidades y deseos. Esto no significa que esfuerzos no científicos como el arte, la literatura y la filosofía hayan perdido valor, o que ciencias más "blandas" como la psicología y la sociología no sean campos importantes del conocimiento. Pero las víctimas del cientificismo opinan lo contrario. Las víctimas del cientificismo eligen creencias que expresan su actitud de exagerado respeto por la ciencia natural. En consecuencia, prefieren creer que las ciencias duras son los únicos campos de conocimiento importantes y

genuinos, y que las cosas estudiadas en las ciencias duras son las únicas que existen. Como analogía aproximada, compárese la probable motivación de la doctrina medieval de que todas las cosas dependen de Dios para su existencia y desaparecerían instantáneamente si Dios dejara de querer que existieran. Los filósofos medievales decían esto, junto con muchas otras cosas inverosímiles, con el propósito de glorificar más a Dios y aumentar su importancia en su filosofía. Del mismo modo, las víctimas del cientificismo, con pocas restricciones, adoptan cualquier creencia filosófica que consideren que glorifica más a la ciencia.

Para una víctima del cientificismo, el intuicionismo ético es obviamente inaceptable, por dos razones. En primer lugar, la ética no forma parte de las ciencias duras. Ni la física, ni la química, ni la biología estudian cosas como la "rectitud". Ni la teoría del big bang, ni la mecánica cuántica, ni la teoría de la evolución pueden explicar por qué el asesinato está mal. Lo máximo que puede hacer la ciencia es explicar las causas de las actitudes morales humanas. Por lo tanto, seguramente, las actitudes son todo lo que hay. Si existieran además hechos morales, entonces habría algo que no estudian los científicos.

En segundo lugar, la intuición, supuestamente, no forma parte del método científico. Creo que esto es bastante falso -pienso que la intuición es esencial para la ciencia[4]- pero se cree ampliamente que es cierto. Las víctimas del cientificismo creen que la ciencia es un conocimiento puramente empírico. Si existiera la intuición, entonces algunos conocimientos serían diferentes del conocimiento científico; por lo tanto, seguramente no existe la intuición. La única forma de realismo moral que las víctimas del cientificismo encuentran digna de consideración es el naturalismo ético, y ello porque los naturalistas éticos tratan de asimilar el conocimiento moral al conocimiento científico. Sólo si se puede demostrar que es "como la ciencia", la ética

puede ser legítima.

9.4.5 Cómo el intuicionismo se volvió "inverosímil"

La fuerte y generalizada oposición al intuicionismo ético, por tanto, es el resultado de la confluencia de al menos tres poderosas y omnipresentes tendencias culturales. La mayoría de los intelectuales modernos han caído presa del cinismo, la corrección política, el cientifismo o alguna combinación de estas actitudes, y ninguna de ellas encaja bien con el intuicionismo. Mi opinión es que esto ha llevado a la caída en desgracia del intuicionismo. Pero no debemos entender este proceso de forma simplista.

En primer lugar, no debemos pensar que los que rechazaron el intuicionismo lo hicieron, por lo general, con un razonamiento explícito como el que he descrito anteriormente; no debemos pensar que las víctimas del cientificismo, por ejemplo, hayan considerado explícitamente el silogismo práctico: "Deseo creer sólo en lo que más glorifica a la ciencia; el intuicionismo ético no glorifica a la ciencia; por lo tanto, rechazaré el intuicionismo". La influencia de actitudes como el cientificismo, el cinismo y la corrección política es más sutil e insidiosa: estas actitudes operan como sesgos, más que como premisas explícitas. Contaminan nuestros juicios sobre lo que parece "plausible" a primera vista, sobre qué teorías merece más la pena investigar, sobre qué objeciones son "importantes" e "interesantes", etc. Así, en lugar de razonar a partir de la premisa "Sólo es verdad lo que glorifica la ciencia", las víctimas del cientifismo simplemente se ven influenciadas a sentir que las teorías que amplían la importancia de la ciencia son "más plausibles" que las que no lo hacen. Las víctimas de lo políticamente correcto simplemente encuentran "inverosímiles" las teorías que implican que las personas sanas son mejores que las otras. Y así

sucesivamente. Por eso, muchos de los que oyen hablar del intuicionismo ético simplemente lo consideran inverosímil a primera vista, por razones que no pueden articular.

Por supuesto, podría haber otras explicaciones: tal vez encuentran el intuicionismo inverosímil porque el intuicionismo está, de hecho, radicalmente equivocada. Pero he dedicado este libro a argumentar que no es así. Es más, si así fuera, esperaríamos que el intuicionismo también le pareciera inverosímil a la gente de muchas otras culturas y en muchas épocas diferentes". En cambio, incluso los antirrealistas parecen admitir que los principios básicos del intuicionismo -en particular el realismo moral- se han asumido a lo largo de la mayor parte de la historia de la filosofía occidental y se presuponen en el discurso moral ordinario.[5] Esto sugiere que cualquiera que considere que el realismo moral es contraintuitivo tiene intuiciones anómalas que exigen alguna explicación especial. Parece más probable que el cambio relativamente reciente en las intuiciones metaéticas que se alejan del realismo moral se deba al tipo de tendencias culturales peculiares que he estado discutiendo, que el hecho de que las intuiciones metaéticas de casi todo el resto de la humanidad a lo largo de la historia hayan estado gravemente sesgadas y que sólo ahora estemos viendo las cosas objetivamente.

A continuación, no debemos suponer que las causas que he mencionado son las únicas responsables del rechazo del intuicionismo, ni debemos suponer que las mismas causas han operado en todas las personas que lo han rechazado. Es muy posible que el cinismo, la corrección política y el cientificismo puedan explicar en última instancia la popularidad del antirrealismo, aunque la mayoría de los antirrealistas no hayan sido ellos mismos víctimas del cinismo, la corrección política o el cientificismo. Para ver por qué esto es así, fíjese en tres puntos: (a) Que una minoría puede constituir una pluralidad; así, si el 40% de las

personas sostienen la teoría *A*, esto puede ser suficiente (siempre que existan más de dos alternativas) para que *A* sea la teoría más extendida. (b) Que las creencias de las personas están fuertemente influenciadas por otras personas; existe un fuerte impulso hacia la conformidad. Por extraño que parezca, probablemente sea cierto que la mayoría de las personas de nuestra sociedad apoyan el antirrealismo sólo porque otras personas lo hacen. (Esas personas probablemente respaldan el antirrealismo a un nivel bastante superficial, y también pueden tener actitudes incompatibles con el antirrealismo). (c) Que algunas personas tienen una influencia intelectual mucho mayor que otras: por ejemplo, un profesor, un periodista o un artista probablemente influyen más en las creencias filosóficas de los demás que un electricista. A los que más influyen en las ideas filosóficas de los demás se les puede llamar "intelectuales". Juntando estos tres puntos, podemos ver que, para que el antirrealismo moral haya llegado a dominar nuestra cultura, podría haber bastado con que una minoría considerable de los intelectuales, que a su vez son una pequeña minoría de la sociedad, hubiera sucumbido al cinismo, a lo políticamente correcto o al cientifismo. El resto de la sociedad puede haber llegado a adoptar el antirrealismo como resultado de la influencia de otros.

No quiero que se entienda que las razones y las pruebas no desempeñan ningún papel en la teorización filosófica. Pero incluso si un individuo acepta sólo aquellas posiciones para las que puede encontrar argumentos plausibles, y las acepta sobre la base de esos argumentos, los prejuicios no racionales pueden seguir desempeñando un papel decisivo en la determinación de su sistema de creencias. Para ver por qué es así, obsérvese (a) que tales prejuicios pueden determinar qué posiciones se buscan argumentos para apoyarlas, (b) que muy a menudo hay argumentos plausibles para cada una de las múltiples teorías que compiten entre sí, especialmente en filosofía, y (c) que los prejuicios no

racionales pueden afectar a lo plausible que uno encuentra un argumento. Teniendo en cuenta los puntos (b) y (c), un filósofo hábil que se ponga a buscar formas de defender una determinada posición filosófica encontrará, por lo general, algunos argumentos que considerará plausibles (naturalmente, esto tiene algunas limitaciones); por lo tanto, el movimiento decisivo es el que se hace justo al principio, cuando uno decide qué teoría investigar con simpatía. Y ese movimiento suele ser más un resultado de la parcialidad que de los argumentos.

Ni que decir tiene que, si estoy en lo cierto, las creencias metaéticas antirrealistas se han formado, en general, de manera poco fiable, mediante procesos que probablemente no conduzcan a la verdad. El hecho de que el antirrealismo permita expresar las actitudes, los sentimientos o la personalidad de uno no es un indicio de su verdad. El hecho de que el realismo moral haga que algunas personas se sientan ofendidas o hiera los sentimientos de la gente no constituye ninguna prueba en su contra. El hecho de que la ciencia sea buena y de que la ciencia no estudie el valor no es una prueba en contra de la realidad del valor.

No espero que mi libro cambie todo esto. Los que se sienten cínicos, los que desean glorificar la ciencia y los que se sienten ofendidos por la idea de que unas culturas tengan más razón que otras seguirán, en su mayoría, rechazando el realismo moral. Si son filósofos, idearán formas ingeniosas de intentar sortear mis argumentos en este libro. Si no son filósofos, ignorarán los argumentos. Sólo unos pocos considerarán seriamente la posibilidad de cambiar sus puntos de vista. Así es, por desgracia, la filosofía.

9.5 La importancia del intuicionismo

No habría escrito este libro si no pensara que la verdad del intuicionismo es importante. Es importante en dos sentidos. En primer lugar, es teóricamente importante. Los intuicionistas proponen un ámbito de hechos -hechos valorativos- diferente de los hechos reconocidos por todos los demás puntos de vista, y una forma de conocer -la intuición- diferente de los reconocidos por todos los demás puntos de vista. La disputa entre los intuicionistas y todos los demás es, pues, una disputa ontológica y epistemológica fundamental. Es crucial para nuestra comprensión general del mundo y nuestro lugar en él.

En segundo lugar, tiene una importancia práctica-moral. Las teorías antirrealistas sobre el valor socavan nuestras creencias morales, nuestra motivación moral e incluso nuestro sentido de la ilación de la vida. Socavan nuestras creencias morales, porque -como admiten los propios antirrealistas- nuestras creencias morales ordinarias y de sentido común son creencias sobre un ámbito putativo de propiedades de valor objetivas. Cuando decimos que algo es moralmente incorrecto, no nos limitamos a decir que va en contra de nuestras preferencias o costumbres, ni pretendemos que nuestro comentario se escuche como una mera expresión de una reacción emocional no cognitiva. La mayoría de la gente entenderá, con razón, que los argumentos antirrealistas implican que todo lo que podemos decir es algo sobre nuestras preferencias o costumbres, o que expresan una actitud no cognitiva, lo que significa que nuestras creencias morales no están justificadas.

Las teorías antirrealistas socavan nuestra motivación moral, porque nos dicen que no tenemos buenas razones para actuar moralmente a menos que hacerlo satisfaga nuestros deseos. En realidad, actuar moralmente a menudo no es, ni nos parece, la mejor manera de satisfacer nuestros deseos. La mayoría de las personas que se convencen del relato humeano de las razones sacarán así la conclusión de que a menudo no

tienen ninguna buena razón para ser morales. (No supondrán, en cambio, que tienen buenas razones para ser morales y concluirán entonces que, en contra de las apariencias, debe ser que el comportamiento moral realmente satisface sus deseos). Además, en la teoría humeana de la motivación, ninguna persona podría estar motivada por otra cosa que no fueran los deseos; incluso si uno fuera un realista moral y pensara erróneamente que ciertos principios morales son objetivamente verdaderos y proporcionan buenas razones para la acción, diría el humeano, todavía no podría estar motivado para actuar simplemente por esas creencias. He argumentado que la teoría humeana de la motivación es falsa, y el apoyo de los filósofos a la teoría no puede hacer que se convierta en verdadera; sin embargo, decir a la gente que necesariamente se comporta de una manera determinada tiene la tendencia a hacer que se comporten más de esa manera de lo que lo harían de otra manera, siempre que crean lo que les dices. Por ejemplo, decir falsamente a la gente que todo el mundo es completamente egoísta tenderá a hacer que -si te creen- empiecen a comportarse de forma más egoísta de lo que lo harían en caso contrario. Y decir falsamente a la gente que nadie puede estar motivado por principios morales objetivos tenderá a hacer que -si te creen- se esfuercen menos por comportarse de acuerdo con los principios morales objetivos. La gente ya tiene suficientes problemas con la debilidad de la voluntad y las tentaciones de racionalizar el comportamiento egoísta, sin que los filósofos les digan que la preocupación moral desinteresada es imposible y proclamen que toda la moral es un "mito", como reza el título de un libro reciente.[6]

Este último punto presupone la veracidad de mi explicación de la motivación moral. Si la explicación humeana de la motivación es correcta, entonces las personas ya actúan únicamente de acuerdo con sus deseos; decirles la teoría humeana no puede hacer que lo hagan más. De hecho, nada de lo que he dicho en esta sección es una prueba para el

intuicionismo, ni es mi razón para apoyar el intuicionismo. Es sólo una razón, dada la verdad del intuicionismo, para pensar que la verdad es importante.

Por último, el antirrealismo moral socava nuestro sentido de la vida, y esto me lleva a una de las razones por las que el antirrealismo me parece increíble. Creo que el antirrealismo se reduce realmente a la opinión de que nada importa. Ahora bien, la mayoría de los antirrealistas lo negarían. Procederían a idear interpretaciones de la frase "nada importa" de manera que signifique algo con lo que su teoría no se compromete. Dirán que "nada importa" expresa una actitud emocional no cognitiva que ellos no tienen de hecho; o que "nada importa" significa que nuestra sociedad no tiene convenciones para aprobar o desaprobar nada; o que significa que ellos mismos no tienen ninguna emoción moral; etc. Todo esto me suena a vacío. A pesar de lo que sé que dirían, creo que el punto de vista antirrealista es, sin embargo, que, en realidad, nada importa. La vida no tiene sentido, porque el "sentido" es una de esas propiedades "espeluznantes", no naturales, en las que los antirrealistas no creen. Los antirrealistas dirían quizás que la vida puede seguir teniendo sentido, porque podemos asignarle algún significado, simplemente adoptando la actitud no cognitiva adecuada hacia ella. Pero esto también suena a vacío. Uno no puede honestamente "asignar" valor a las cosas mientras considera conscientemente que esa asignación es meramente arbitraria, y mantener que desde el punto de vista objetivo, esas cosas no tienen valor. Los valores morales, como todas las creencias, son cosas celosas: contienen una pretensión inherente de corrección. Si creo que *p*, estoy creyendo que *p* es verdadera; no puedo reconocer sistemáticamente la negación de *p* como una creencia igualmente aceptable, o igualmente verdadera. Del mismo modo, si creo sinceramente que *x* es bueno, considero que mi juicio de valor es correcto; no puedo reconocer coherentemente una perspectiva

igualmente (o más) correcta que no asigne ningún valor a *x*.

Hay que admitir que es posible -como han hecho los subjetivistas y los no cognitivistas- idear interpretaciones del discurso evaluativo de manera que se pueda negar el valor objetivo, y sin embargo, decir que las cosas son "realmente buenas" o "realmente malas". Pero, como he argumentado en capítulos anteriores, estas interpretaciones son simplemente falsas. Cuando digo "*p* es verdadero", no quiero decir simplemente "creo que *p*", y esto, contra los relativistas de la verdad, es algo sobre lo que no hay ninguna duda. Análogamente, cuando digo "*x* es bueno", no quiero decir simplemente "valoro *x*", o algo similar. En resumen, creo que entre los antirrealistas, sólo los nihilistas son totalmente francos.

9.6 Cómo me convertí en un intuicionista espeluznante y anticientífico

He sido un realista moral desde que tengo uso de razón. Creo que la razón es más o menos ésta: me parece que ciertas cosas, como el dolor y el sufrimiento, por poner el ejemplo más claro, son malas. No creo que me lo esté inventando, y no creo que sea sólo una preferencia personal arbitraria mía. Si pongo el dedo en una llama, tengo una determinada experiencia, y puedo ver directamente algo en ella (en la experiencia) que es malo.[7] Además, si es malo que yo experimente dolor, parece que también debe ser malo que otra persona experimente dolor. Por lo tanto, no debo infligir ese dolor a los demás, como tampoco ellos deben infligirlo a mí.[8] Así que hay al menos un ejemplo de un principio moral racional.

Pero no siempre he sido un intuicionista: cuando conocí la teoría, en la forma defendida por W. D. Ross, pensé que hacía que la ética fuera

inaceptablemente arbitraria. Ross había dado una lista de diferentes "deberes *prima facie*" que tenemos, como el deber de cumplir las promesas, el deber de evitar causar daño, el deber de mostrar gratitud por los beneficios que nos han dado, etc. Ross no explicaba por qué todas estas cosas eran deberes, aparte de que a la mayoría de la gente le parecía que lo eran; tampoco explicaba cómo resolver los conflictos entre los deberes prima facie en casos particulares, aparte de usar el mejor juicio de cada uno.[9] Me parecía que esto hacía que la mayoría de los juicios morales fueran doblemente arbitrarios: había arbitrariedad tanto en los principios generales del deber como en su aplicación a circunstancias particulares.

No estoy seguro, ahora, de por qué pensaba estas cosas. Si alguien me hubiera preguntado: "¿Cree usted que toda creencia requiere una serie infinita de argumentos para estar justificada?", seguramente habría dicho que no. Así que lo que era "arbitrario" en las opiniones éticas de Ross no podía ser el mero hecho de que presentara algunas afirmaciones sin argumentos. Creo que me habría sentido satisfecho, o al menos, menos inclinado a hacer la objeción de "arbitrariedad", si Ross hubiera enunciado un único principio ético del que pudieran derivarse lógicamente todos los demás principios éticos, como hacen, por ejemplo, los utilitaristas. Nadie acusa de arbitrariedad a los utilitaristas. Pero esto tiene poco sentido: si un único principio ético fundacional puede ser no arbitrario, ¿por qué no dos? ¿O seis? ¿O cien?

En cualquier caso, no fueron las consideraciones anteriores las que me llevaron al punto de vista de Ross. Por el contrario, el punto de inflexión se produjo cuando me di cuenta de que había muchas cosas que sabía que nunca podría demostrar de ninguna manera científica e intersubjetivamente accesible. Y no se trataba de una especie de revelaciones místicas, sino de cosas perfectamente mundanas de las que

no había ninguna duda seria de que las sabía. Un amigo podía llamarme por teléfono y, después de tal vez un segundo de escucharlo, sabría de quién se trataba. No tendría que decir su nombre ni ningún otro dato que la identificara. Si alguien me preguntaba cómo sabía de quién se trataba, no podía decir nada más informativo que: "Bueno, suena como mi amigo". Por supuesto, debe haber alguna combinación de características del patrón de voz que me permita reconocerla. Quizá los psicólogos cognitivos podrían averiguarlo. Pero, desde luego, yo no podría decírselo. Sin embargo, no hay duda alguna de que conozco la voz de mi amigo.

Otros ejemplos de conocimiento indemostrable son fáciles de conseguir. Hablando con un amigo, me doy cuenta de que está disgustado por algo; no podría decirle cómo lo sé. No podría darle una definición satisfactoria de "mesa", pero sé que el objeto que tengo delante es una, y el objeto del que estoy bebiendo no lo es. Los filósofos modernos han tendido a ver la ciencia y las matemáticas como los paradigmas del conocimiento, y los casos de conocimiento indemostrable como excepciones periféricas a la regla. Pero me di cuenta de que esto estaba muy sesgado: el conocimiento indemostrable es la norma. Esto no quiere decir nada en contra del conocimiento científico o matemático, pero son las excepciones: una persona normal puede vivir toda su vida y no hacer nunca una prueba matemática o un experimento científico. Pero cada uno de nosotros llega a conocer constantemente el tipo de cosas que aparecen en mis ejemplos anteriores.

Así que la confianza de Ross en los ejercicios de juicio intelectual en casos individuales, en ausencia de cualquier algoritmo identificable, no hace que su visión del conocimiento moral sea extraña, en comparación con la mayoría de la cognición humana. Cuando consideramos ciertas situaciones, como el problema del tranvía o el ejemplo de la cosecha de órganos,[10] encontramos que podemos juzgar ciertas acciones como

correctas o incorrectas, aunque carezcamos de cualquier teoría sistemática capaz de computar la respuesta correcta en todos los casos. Esto no hace que nuestros juicios sean "arbitrarios", al igual que no es arbitraria mi creencia de que estoy hablando con mi amigo por teléfono. Se necesitaría una disposición absurdamente escéptica para declarar que, como no puedo especificar un conjunto de propiedades del patrón de voz por el que se pueda identificar la voz de mi amigo, no sé con quién estoy hablando.

Un segundo punto de inflexión importante se produjo cuando asistí a una conferencia de Donald Davidson en Berkeley explicando su "monismo anómalo". La conferencia de Davidson se refería a la relación entre los estados mentales y los estados cerebrales, pero una visión análoga podría tomarse de la relación entre las propiedades morales y las propiedades naturales. Davidson opinaba que los estados mentales superponen a los estados cerebrales -es imposible que se produzca un cambio en el estado mental de una persona sin que se produzca algún cambio en el cerebro- pero que, sin embargo, no es posible seleccionar todos los estados mentales de un tipo determinado utilizando únicamente conceptos de la fisiología cerebral. Por ejemplo, el estado mental de creer que los mosquitos tienen dientes pequeños podría ser realizado por muchos estados cerebrales diferentes, que podrían no tener nada interesante en común cuando se describen únicamente en términos de fisiología cerebral. Davidson explicó extensamente que no había ninguna dificultad conceptual en esto. Para mí, lo importante era que el hecho evidente de que las propiedades morales dependen de las propiedades naturales no implica que se pueda identificar alguna propiedad natural, o incluso una disyunción finita de propiedades naturales, que tengan en común todas las acciones moralmente correctas. Por lo tanto, no había nada metafísicamente absurdo en la imagen de Ross de múltiples deberes que no tienen ninguna característica natural en común, que no

tienen nada en común más que el hecho de ser deberes.

Con ello, mi principal fuente de resistencia al intuicionismo -la objeción de la "arbitrariedad"- quedó desactivada, y fui libre de aceptar que, en el ámbito moral, las cosas son, después de todo, esencialmente como parecen. Tengo intuiciones morales, como casi todo el mundo. Parecen ser cogniciones de verdades morales. Parece que cosas como el asesinato, la tortura y el robo están realmente mal, y no sólo porque a mí personalmente no me gusten. No pretendo ser infalible, ni para mí ni para la comunidad humana; es lógicamente posible que yo y el resto de la humanidad estemos radicalmente engañados (también es "posible" en el mismo sentido que todos somos cerebros en cubas). Pero ¿no deberíamos ver primero si se puede mantener que todo es más o menos como parece, antes de saltar a teorías esquizoides sobre engaños y errores radicales? Una teoría plausible de la moral no puede ignorar sin más nuestras intuiciones morales. ¿No sería más sencillo decir que la razón por la que tenemos la intuición de que el dolor es malo es porque el dolor es, de hecho, malo, y nuestra intuición es nuestra conciencia de ese hecho?

Notas

1. Esta objeción supone que el subjetivista apoya el reduccionismo analítico. Si la identidad entre lo malo y lo prohibido por Dios es sintética, entonces sería coherente pensar que "robar es malo, pero Dios no lo prohíbe". Del mismo modo, si la identidad entre lo incorrecto y lo prohibido por la sociedad es sintética, entonces se podría pensar de forma coherente que la poligamia es permisible pero la sociedad la prohíbe, por lo que dos personas podrían estar en desacuerdo sobre la permisividad de la poligamia mientras están de acuerdo en que la sociedad la prohíbe. Una idea similar se aplica al subjetivismo individual.

2. Este argumento asume el principio de bivalencia (toda proposición es verdadera o falsa) y la teoría de la correspondencia de la verdad (la verdad es la correspondencia con

los hechos). Hay importantes objeciones a la bivalencia, como las derivadas de la vaguedad y de los nombres vacíos, pero ninguna de ellas es relevante aquí.

3. Van Inwagen (1983, p. 215).

4. Véase BonJour (1998, pp. 203-6) para una explicación parcial de por qué esto es así.

5. Véase Mackie (1977, pp. 30-5).

6. Joyce (2001).

7. Véase Stuart Rachels (2003, pp. 17-22) para profundizar en este punto.

8. El paso a la obligación de no dañar a otros es controvertido. Véase Parfit (1984, capítulos 6, 7 y 9) y Nagel (1970) para defensas más sofisticadas de este tipo de movimiento. Pero el simple pensamiento del texto es el que me mueve.

9. Ross ([1930]1988, pp. 21, 30-1).

10. Véase más arriba, p. 103.

11. Lo que sigue es una versión simplificada de una parte de las opiniones de Davidson sobre lo mental; véase Davidson (1991).

Referencias

Abraham, Leo. 1933. 'The Logic of Ethical Intuitionism'. International Journal of Ethics 44:37-55.

Adams, Robert Merrihew. 1981. 'A Modified Divine Command Theory of Ethical Wrongness'. Pp. 83-108 in Divine Commands and Morality, edited by Paul Helm. Oxford: Oxford University Press.

---. 1999. Finite and Infinite Goods. New York: Oxford University Press.

Allen, Elizabeth, et al. 1977. 'Sociobiology: A New Biological Determinism'. In Biology as a Social Weapon, edited by Sociobiology Study Group of Boston. Minneapolis, Minn.: Burgess Publishing Co.

Alston, William. 1976. 'Has Foundationalism Been Refuted?' Philosophical Studies 29:287-305.

---. 1990. 'Some Suggestions for Divine Command Theorists'.

In Christian Theism and the Problems of Philosophy, edited by Michael Beaty. Notre Dame, Ind.: University of Notre Dame Press.

---. 1993. The Reliability of Sense Perception. Ithaca, N. Y.: Cornell

University Press.

Anscombe, G. E. M. 1963. Intention. 2nd ed. Ithaca, N.Y.: Cornell University Press.

Arens, William. 1979. The Mall-Eating Myth: Anthropology and Allthropophagy. New York: Oxford University Press.

Armstrong, David. 1978a. Nominalism and Realism. Cambridge:

Cambridge University Press.

---. 1978b. A Theory of Universals. Cambridge: Cambridge University Press.

Audi, Robert. 1983. 'Foundationalism and Epistemic Dependence'.

Synthese 55:119-39.

---. 1997. Moral Knowledge and Ethical Character. New York:

Oxford University Press.

---. 2004. The Good in the Right. Princeton, N.J.: Princeton University Press.

Austin,J. L. 1975. How to Do Things with Words. 2nd ed. Edited by J. o.

Urmson and Marina Sbisa. Cambridge, Mass.: Harvard University Press.

Ayer, A. J. 1952. Language, Truth, alld Logic. New York: Dover.

---. 1956. The Problem of Knowledge. London: Macmillan.

Bealer, George. 1992. 'The Incoherence of Empiricism'. Proceedings of the Aristotelian Society 66, supplement: 99-138.

---. 1993. 'Universals'. Journal of Philosophy 40:5-32.

---. 1999. 'A Theory of the A Priori'. Pp. 29-55 in Philosophical

Perspectives 13: Epistemology, edited by James Tomberlin. Cambridge, Mass.: Blackwell.

Benacerraf, Paul. 1973. 'Mathematical Truth'. Journal of Philosophy 70:661-79.

Benedict, Ruth. 1934. 'Anthropology and the Abnormal'. Journal of General Psychology 10:59-82.

Bentham, Jeremy. [1781] 1996. An Introduction to the Principles of Morals and Legislation. Edited by]. H. Burns and H. L. A. Hart.

Reprint, Oxford: Clarendon Press.

Berkeley, George. [1710] 1965. Treatise COllcerning the Principles of Human Knowledge. Reprinted in Prillciples, Dialogues, and Correspondence, edited by Colin Turbayne. Indianapolis, Ind.: Bobbs-Merrill.

Blackburn, Simon. 1984. Spreading the Word. Oxford: Oxford University Press.

---. 1993. Essays in Quasi-Realism. Oxford: Oxford University Press.

---.1998. Ruling Passions: A Theory of Practical Reasoning. Oxford: Clarendon Press.

Bonjour, Laurence. 1985. The Structure of Empirical Knowledge.

Cambridge, Mass.: Harvard University Press.

---. 1998. In Defense of Pure Reason. Cambridge: Cambridge University Press.

---. 2004. 'In Search of Direct Realism'. Philosophy and Phenomenological Research 69:349-67.

---. 2005. 'In Defense of the a Priori', 'Reply to Devitt', and 'Last Rejoinder'. Pp. 98-122 in COlltemporary Debates in Epistemology,

edited by Matthias Steup and Ernest Sosa. Malden, Mass.: Blackwell.

Borowski, E. J. 1980. 'Moral Autonomy Fights Back'. Pllilosophy 55:95-100.

Boyd, Richard. 1988. 'How to Be a Moral Realist'. Pp. 181-229 in Essays on Moral Realism, edited by Geoffrey Sayre-McCord. Ithaca,

N.Y.: Cornell University Press.

Brandt, Richard. 1944. 'The Significance of Differences of Ethical Opinion for Ethical Rationalism'. Philosophy and Phenomenological Research 4:469-94.

---.1979. A Theory of the Good alld the Right. Oxford: Clarendon.

Brandt, Richard. 1996. Facts, Values, and Morality. Cambridge: Cambridge University Press.

Brighouse, M. H. 1990. 'Blackburn's Projectivism-An Objection'.

Philosophical Studies 59:225-33.

Brink, David. 1984. 'Moral Realism and the Sceptical Arguments from Disagreement and Queerness'. Australasian Journal of Philosophy 62:111-25.

Butchvarov, Panayot. 1989. Skepticism in Ethics. Bloomington, Ind.: Indiana

University Press.

Butler, joseph. [17 29J 1964. Butler's Sermons and Dissertation on Virtue.

London: G. Bell & Sons.

Carson, Thomas. 2000. Value and the Good Life. Notre Dame, Ind.: University of Notre Dame Press.

Chisholm, Roderick. 1982. 'The Problem of the Criterion'. Pp. 61-75 in The Foundations of Knowing. Minneapolis, Minn.: University of Minnesota Press.

Clarke, Samuel. [1705J 1964. 'On Natural Religion'. Reprinted in British Moralists, vol. 2, edited by L. A. Selby-Bigge. Indianapolis, Ind.: Bobbs-Merrill.

Crane, Hewitt and Thomas Piantanida. 1983. 'On Seeing Reddish Green and Yellowish Blue'. Science 221:1078-80.

Dancy, jonathan. 1993. Moral Reasons. Oxford: Blackwell.

Darwall, Stephen, Allan Gibbard, and Peter Railton. 1992. 'Toward Fin de Siecle Ethics: Some Trends'. Philosophical Review 101: 115-89.

Davidson, Donald. 1990. 'A Coherence Theory of Truth and Knowledge'. In Reading Rorty, edited by Alan Malachowski. Cambridge, Mass.: Basil Blackwell.

---. 1991. 'Mental Events'. Pp. 247-56 in The Nature of Mind, edited by David M. Rosenthal. New York: Oxford University Press.

Dawkins, Richard. 1989. The Selfish Gene. Oxford: Oxford University Press.

DePaul, Michael. 1988. 'The Problem of the Criterion and Coherence Methods in Ethics'. Canadialllournal of Philosophy 18:67-86.

Dreier, james. 1996. 'Expressivist Embeddings and Minimalist Truth'.

Philosophical Studies 83:29-51.

---. 2002. 'Troubling Developments in Metaethics'. Nous 36:152-68.

Firth, Roderick. 1952. 'Ethical Absolutism and the Ideal Observer'.

Philosophy and Phenomellological Research 12:317-45.

Foley, Richard. 1979. 'justified Inconsistent Beliefs'. American Philosophical Quarterly 16:247-57.

Foley, Richard and Richard Fumerton. 1985. 'Davidson's Theism?'

Philosophical Studies 48:83-9.

Foot, Philippa. 1967. 'The Problem of Abortion and the Doctrine of the Double Effect'. Oxford Review 5:5-15.

Frankena, William. 1958. 'Obligation and Motivation in Recent Moral Philosophy'. In Essays in Moral Philosophy, edited by A. I. Melden.

Seattle: University of Washington Press.

---. 1973. Ethics. Englewood Cliffs, N.J.: Prentice-Hall.

Freeman, Derek. 1983. Margaret Mead and Samoa: The Making and Unmaking of an Anthropological Myth. Cambridge, Mass.: Harvard University Press.

---. 1999. The Fateful Hoaxing of Margaret Mead: A Historical Analysis of Her Samoan Research. Boulder, Colo.: Westview.

Friedman, David. 1994. 'A Positive Account of Property Rights'. Social Philosophy and Policy 11:1-16.

Fumerton, Richard. 1990. Reason and Morality: A Defense of the Egocentric Perspective. Ithaca, :"-J. Y.: Cornell University Press.

---. 1995. Metaepistemology and Skepticism. Lanham, Md.:

Rowman & Littlefield.

---. Forthcoming. Epistemology. Blackwell.

Gallup Organization. n.d. 'Moral Issues'. (Most recent poll results dated May, 2005.) <http://www.gallup.com/poll/content/default. aspx?ci=1681>. Accessed 26 May 2005.

Gallup, George H., Jr., and Frank Newport. 1991. 'Belief in Paranormal

Phenomena Among Adult Americans'. Skeptical Inquirer 15,2: 137-46.

Gardner, Martin. 1998. 'Is Cannibalism a Myth?' Skeptical Inquirer 22, 1: 14-16.

Gay, Robert. 1985. 'Ethical Pluralism: A Reply to Dancy'. Minci 94:250-62.

Geach, P. T. 1965. 'Assertion'. Philosophical Review 74:449-65.

---. 1977. 'Again the Logic of "Ought"'. Philosophy 52:473-6.

Gettier, Edmund. 1963. 'Is Justified, True Belief Know ledge?' Analysis 23:121-3.

Gibbard, Allan. 1990. Wise Choices, Apt Feelings. Cambridge, Mass.: Harvard University Press.

Gilovich, Thomas. 1993. How We Know What Isn't So. New York: Free Press.

Glassen, Peter. 1959. 'The Cognitivity of Moral Judgments'. Mind 68:57-72.

Goldhagen, Daniel. 1996. Hitler's Willing Executioners: Ordinary Germans and the Holocaust. New York: Alfred A. Knopf.

Goldman, Alvin. 1967. 'A Causal Theory of Knowing'. Journal of Philosophy 64:357-72.

Haack, Susan. 1983. 'Theories of Knowledge: An Analytic Framework'.

Proceedings of the Aristotelian Society 83:143-57.

Haidt, Jonathan. 2001. 'The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment'. Psychological Review 108:814-34.

Haidt, Jonathan and Matthew A. Hersh. 2001. 'Sexual Morality: The Cultures and Emotions of Conservatives and Liberals'. Journal of Applied Psychology 31:191-221.

Haidt, Jonathan, Silvia Helena Koller, and Maria G. Dias. 1993.

'Affect, Culture, and Morality, or Is It Wrong to Eat Your Dog?'

Journal of Personality and Social Psychology 65:613-28.

Hare, R. M.1952. The Language of Morals. Oxford: Oxford University Press.

---. 1963. Freedom and Reasoll. Oxford: Clarendon Press.

---. 1970. 'Meaning and Speech Acts'. Philosophical Review 79:3-24.

---.1989. 'Rawls' Theory ofjustice'. In Reading Rawls, edited by Norman Daniels. Stanford, Calif.: Stanford University Press.

---. 1999. Objective Prescriptions and Other Essays. Oxford: Clarendon.

Harman, Gilbert. 1977. The Nature of Morality: An Introduction to Ethics. New York: Oxford University Press.

---. 1990. 'The Intrinsic Quality of Experience'. Pp. 31-52 in Philosophical Perspectives 4: Action Theory and Philosophy of Mind, edited by James Tomberlin. Atascadero, Calif.: Ridgeview Publishing Co.

Hellman, Hal. 1998. Great Feuds in Science. New York: John Wiley & Sons.

Hobbes, Thomas. [1651] 1996. Leviathan. Edited by Richard Tuck.

Reprint, Cambridge: Cambridge University Press.

Horgan, Terence and Mark Timmons. 1992. 'Troubles for New Wave Moral Semantics: The "Open Question Argument" Revived'.

Philosophical Papers 21:153-75.

Howard, Michael and Janet Dunaif-Hattis. 1992. Anthropology: Understanding Human Adaptation. New York: HarperCollins.

Hubin, D. Clayton. 1980. 'Prudential Reasons'. Canadiall Journal of Philosophy 10:63-81.

Hudson, W. D. 1967. Ethical Intuitionism. New York: St. Martin's Press.

Hudson, W. D. 1983. Modern Moral Philosophy. 2 nd ed. New York: St.

Martin's Press.

Huemer, Michael. 1997. 'Probability and Coherence Justification'.

Southern Journal of Philosophy 35:463-72.

---. 1999. 'The Problem of Memory Knowledge'. Pacific Philosophical Quarterly 80:346-57.

---. 2000a. 'Naturalism and the Problem of Moral Knowledge'. Southern Joumal of Philosophy 38:575-97.

---. 2000b. 'Van Inwagen's Consequence Argument'. Philosophical Review 109:524-43.

---. 200l. Skepticism and the Veil of Perception. Lanham, Md.: Rowman & Littlefield.

---.2003. 'Arbitrary Foundations?' Philosophical Forum 34:141-52.

---. n.d. 'Why People Are Irrational about Politics'. <http://home.sprynet.com/-owll/irrationality.htm>. Accessed 28 May 2005.

---. Forthcoming. 'Compassionate Phenomenal Conservatism'.

Philosophy and Phenomenological Research.

Hume, David. [l75 1]1975. Enquiry Concerning the Principles of Morals.

Reprinted in Enquiries Concerning Human Understanding and Concerning the Principles arM orals, edited by P. H. Nidditch. Oxford: Clarendon Press.

---. [1758] 1975. Enquiry Concerning Human Understanding.

Reprinted in Enquiries Concerning Human Understanding and Concerning the Principles ofM orals, edited by P. H. Nidditch. Oxford: Clarendon Press.

---. [1739] 1992. Treatise of Human Nature. Reprint, Buffalo, N.Y.: Prometheus.

Hurka, Thomas. 1980. 'Geach on Deriving Categorical "Oughts"'.

Philosophy 55:101-4.

Jackson, Frank. 2003. 'Cognitivism, A Priori Deduction, and Moore'.

Ethics 113:557-75.

Jackson, Frank, Graham Oppy, and Michael Smith. 2004. 'Minimalism and Truth

Aptness'. Pp. 233-51 in Mind, Morality, and Explanation, edited by Frank Jackson, Philip Pettit, and Michael Smith. Oxford: Clarendon.

Jackson, Frank, Philip Pettit, and Michael Smith. 2004. 'Ethical Particularism and Patterns'. Pp. 211-32 in Mind, Morality, and Explanation, edited by Frank Jackson, Philip Pettit, and Michael Smith. Oxford: Clarendon.

Joyce, Richard. 2001. The Myth of Morality. Cambridge: Cambridge University Press.

Kant, Immanuel. [1785] 1959. Founddtions of the Metaphysics of Morals. Translated by Lewis White Beck. Reprint, Indianapolis, Ind.: Bobbs-Merrill.

Karmo, Toomas. 1988. 'Some Valid (but no Sound) Arguments Trivially Span the IIs"-"0ught" Gap'. Mind 97:252-7.

Kim, Jaegwon. 1993. 'Concepts of Supervenience'. Pp. 53-78 in Supervenience and Mind. Cambridge: Cambridge University Press.

Kripke, Saul. 1972. Naming and Necessity. Cambridge, Mass.: Harvard University Press.

Layman, C. Stephen. 1991. The Shape of the Good: Christian Reflections on the Foundations of Ethics. Notre Dame, Ind.: University of Notre Dame Press.

LaughLab. n.d. <http://www.Iaughlab.co.uk/>. Accessed 27 May 2005.

Lemos, Noah. 1994. Intrinsic Value. Cambridge: Cambridge University Press.

Lewis, David. 1986. On the Plurality of Worlds. New York: Blackwell.

Liszka, James Jakob. 2002. Moral Competence: An Integrated Approach to the Study of Ethics . 2nd ed. Upper Saddle River, N.].: Prentice Hall.

Mackie, John L. 1977. Ethics: Inventing Right and Wrong. New York: Penguin.

Maitzen, Stephen. 1998. 'Closing the IIs"-"Ought" Gap'. Canadian Journal of Philosophy 28:349-66.

Marquis, Donald. 1989. 'Why Abortion Is Immoral'. Journal of Philosophy 86: 183-202.

McDowelI, John. 1978. 'Are Moral Requirements Hypothetical Imperatives?' Proceedings of the Aristotelian Society 52, supplement: 13-29.

McGrath, Sarah. 2004. 'Moral Knowledge by Perception'. Pp. 209-28 in Philosophical Perspectives 18: Ethics, edited by John Hawthorne.

Malden, Mass.: Blackwell.

McNaughton, David. 1988. Moral Vision: An Introduction to Ethics.

Oxford: Blackwell.

---. 2000. 'Intuitionism'. Pp. 268-87 in The Blackwell Guide to Ethical Theory, edited by Hugh LaFollette. Oxford: Blackwell.

Mead, Margaret. 2001. Coming of Age in Samoa. New York: HarperCollins.

Moore, G. E. 1947. Ethics. London: Oxford University Press.

---.1953. 'Hume's Theory Examined'. Pp. 108-26 in Some Main Problems of Philosophy. London: Allen & Unwin.

1960. Principia Ethira. Cambridge: Cambridge University Press.

Moore, Michael S. 1992. 'Moral Reality Revisited'. Michigan Law Review 90:2424-533.

Murray, Charles. 1984. Losing Ground: American Social Policy, 1950-1980. New York: Basic Books.

Nagel, Ernest and James R. Newman. 2001. Godel's Proof Revised ed. New York: New York University Press.

Nagel, Thomas. 1986. The View from Nowhere. New York: Oxford University Press.

---. 1970. The Possibility of Altruism. Princeton, N.J.: Princeton University Press.

Nelson, Mark. 1991. 'Intuitionism and Subjectivism'. Metaphi/osoplly 22:115-21.

---. 2003. 'Sinnott-Armstrong's Moral Scepticism'. Ratio 16: 63-82.

Newport, Frank and Maura Strausberg. 2001. 'Americans' Belief in Psychic and Paranormal Phenomena Is up Over Last Decade'. The Gallup Organization. June 8. <http://www.gallup.com/poll/content/?ci=4483&pg=1>. Accessed 26 May 2005.

Nowell-Smith, P. H. 1954. Ethics. Baltimore, Md.: Penguin.

Nozick, Robert. 1974. Anarchy, State, and Utopia. New York: Basic Books.

---.1981. Philosophical Explanations. Cam bridge, Mass.: Harvard University Press.

Olsson, Erik. 2005. Against Coherence: Truth, Probability, and Justification. Oxford: Clarendon Press.

Parfit, Derek. 1984. Reasons and Persons. Oxford: Clarendon.

Pigden, Charles. 1989. 'Logic and the Autonomy of Ethics'. Australasian Journal of Philosophy 67: 127-5 l.

---. 1991. 'Naturalism'. Pp. 421-31 in A Companion to Ethics, edited by Peter Singer. Oxford: Basil Blackwell.

Pinker, Steven. 2002. The Blank Slate: The Modem Denial of Human Nature. New York: Penguin.

Plantinga, Alvin. 1993. Warrallt: Tile Currellt Debate. Oxford: Oxford University Press.

Plato. 1996. Plato's Euthyphro, Apology, and Crito. Edited by S. W.

Emery. Lanham, Md.: University Press of America.

Popper, Karl. 2001. 'Facts, Standards, and Truth: A Further Criticism of Relativism'. Pp. 32-52 in Moral Relativism, edited by Paul Moser and Thomas Carson. Oxford: Oxford University Press.

Prichard, H. A. 1957. Moral Obligatioll. Oxford: Clarendon Press.

Prior, A. N. 1976. 'The Autonomy of Ethics'. Pp. 88-96 in Papers in Logic and Ethics, edited by P. T. Geach and A. J. P. Kenny. London: Duckworth.

Putnam, Hilary. 1975. 'Philosophy and our Mental Life'. Pp. 291-303 in Mind, Language and Reality. Vol. 2 of Philosophical Papers.

Cambridge: Cambridge University Press.

---. 1981. Reason, Truth and History. Cambridge: Cambridge University Press.

Quine, W. V. 1976. 'Truth by Convention'. Pp. 77-106 in The Ways of Paradox. Revised ed. Cambridge, Mass.: Harvard University Press.

Quinn, Philip. 2002. 'God and Morality'. In Reason and Responsibility, 11th ed., edited by Joel Feinberg and Russ Shafer-Landau. Belmont, Calif.: Wadsworth.

Quinn, Warren. 1993. 'Putting Rationality in Its Place'. Pp. 228-55 in Morality and Action. Cambridge: Cambridge University Press.

Rachels, James. 1974. 'Two Arguments against Ethical Egoism'.

Philosophia 4:297-314.

---. 1990. Created from Animals: The Moral Implications of Darwinism. Oxford: Oxford University Press.

---. 2003. The Elements of Moral Philosophy, 4th ed. Boston: McG raw-Hill.

Rachels, Stuart. 1998. 'Counterexamples to the Transitivity of Better Than'. Australasian Journal of Philosophy 76:71-83.

---. 2003. 'A Defense of Two Optimistic Claims in Ethical Theory'. Philosophical Studies 112:1-30.

---. 2004. 'Six Theses About Pleasure'. Pp. 247-67 in Philosophical Perspectives 18: Ethics, edited by John Hawthorne. Malden, Mass.: Blackwell.

Railton, Peter. 1986. 'M oral Realism'. Philosophical Review 95: 163-207.

---.1998. 'Moral Explanation and Moral Objectivity'. Philosophy and Phenomenological Research 58: 175-82.

Rand, Ayn. 1964. The Virtue of Selfishness. New York: Signet.

---. 1990. Introduction to Objectivist Epistemology. 2 nd ed. New York: NAL

Books.

Rashdall, Hastings. 1907. The Theory of Good and Evil: A Treatise on Moral Philosophy. Oxford: Clarendon.

Rawls, John. 1971. A Theory of Justice. Cambridge, Mass.: Harvard University Press.

Reid, Thomas. [17881 1983. Essays on the Active Powers of Man.

Selections reprinted in Inquiry and Essays, edited by Ronald Beanblossom and Keith Lehrer. Indianapolis, Ind.: Hackett.

Ross, W. D. [193011988. The Right and tile Good. Reprint, Indianapolis, Ind.: Hackett.

Ruse, Michael. 1985. Sociobiology: Sense or Nonsense? Dordrecht: D. Reidel.

Ruse, Michael. 1988. Homosexuality: A Philosophical Inquiry. Oxford: Basil Blackwell.

---. 1998. Taking Darwin Seriously. Amherst, N.Y.: Prometheus.

Russell, Bertrand. [19121 1997. The Problems of Philosophy. Reprint, Oxford: Oxford University Press.

Ryan, Sharon. 2003. 'Doxastic Compatibilism and the Ethics of Belief'. Philosophical Studies 114:47-79.

Scanlon, T. M. 1998. What We Owe to Each Other. Cambridge, Mass.: Harvard University Press.

Searle, John R. 1964. 'How to Derive "Ough til from "ls"'. Philosophical Review 73:43-58.

---. 1969. Speech Acts. Cambridge: Cambridge University Press.

---.1979. Expression and Meaning: Studies in the Theory of Speech Acts. Cambridge: Cambridge University Press.

---. 1983. Intentionality: An Essay in the Philosophy of Mind. Cambridge: Cambridge University Press.

Sextus Empiricus. 1964. Scepticism, Man, and God. Edited by Philip Hallie and translated by Sanford Etheridge. Middletown, Conn.: Wesleyan University Press.

Shafer-Landau, Russ. 2003. Moral Realism. Oxford: Clarendon.

Shandley, Robert R., ed. 1998. Unwilling Germans? The Goldhagen Debate. Minneapolis, Minn.: University of Minnesota Press.

Singer, Peter. 1974. 'Sidgwick and Reflective Equilibrium'. The Mon ist 58:490-517.

---. 1993. Practical Ethics, 2nd ed. Cambridge: Cambridge University Press.

---. 1999. 'A Response'. Pp. 269-335 in Singer and His Critics, edited by Dale Jamieson. Oxford: Blackwell.

Singer, Peter and Richard Posner. 2001. 'Animal Rights: Debate between Peter Singer and Richard Posner'. Slate Magazine, June 11-15 <**Error! Hyperlink reference not valid.**>.

Sinnott-Armstrong, Waiter. 1996. 'Moral Skepticism and Justification'. Pp. 3-48 in Moral Knowledge? New Readings in Moral Epistemology, edited by Walter Sinnott-Armstrong and Mark Timmons.

Oxford: Oxford University Press.

Smart, J. J. C. 1973. 'An Outline of a System of Utilitarian Ethics'. In Utilitarianism: For and Against by j. j. C. Smart and Bernard Williams. Cambridge: Cambridge University Press.

Smith, Michael. 1994. The Moral Problem. Oxford: Basil Blackwell.

Smith, Tara. 2000. Viable Values. Lanham, Md.: Rowman & littlefield.

Sterling, Grant. 1994. Ethical Intuitionism and Its Critics. New York:

Peter Lang.

Stevenson, Charles L. 1963. Facts and Values: Studies in Ethical Analysis. New Haven, Conn.: Yale University Press.

Stratton-Lake, Philip. 2002a. Introduction to The Right and the Good, by W. D. Ross. Oxford: Oxford University Press.

---, ed. 2002b. Ethical Intuitionism: Re-evaluations. Oxford: Oxford University Press.

Strawson, P. F. 1949. 'Ethical Intuitionism'. Philosophy 24:23-33.

Sturgeon, Nicholas. 1985. 'Moral Explanations'. In Morality, Reason and Truth, edited by David Copp and David Zimmerman. Totowa, N.J.: Rowman & Allanheld.

---. 1986. 'Harman on Moral Explanations of Natural Facts'.

Southern Journal of Philosophy 24, supplement: 69-78.

---. 1991. 'Contents and Causes: A Reply to Blackburn'.

Philosophical Studies 61:19-37.

Sumner, William Graham. 1907. Folkways. New York: Ginn and Company.

Thompson, Bert. 1999. 'The Bible and the Age of the Earth'. Reason and Revelation 19,8: 57-63.

Thomson,Judith Jarvis. 1971. 'A Defense of Abortion'. Philosophy and Public Affairs 1:47-66.

Timmons, Mark. 1999. Morality without Foundations: A Defense of Ethical Contextualism. New York: Oxford University Press.

Tooley, Michael. 1972. 'Abortion and Infanticide'. Philosophy and Public Affairs 2:37-65.

Tye, Michael. 2002. 'Representationalism and the Transparency of Experience'. Nous 36:137-51.

Unger, Peter. 1996. Living High and Letting Die: Our Illusion of Innocence. New York: Oxford University Press.

Vallentyne, Peter. 1996. 'Response-Dependence, Rigidification and Objectivity'. Erkenntllis 44:101-12.

Van Inwagen, Peter. 1983. An Essay on Free Will. Oxford: Clarendon Press.

Van Roojen, Mark. 1996. 'Expressivism and Irrationality'. Pililosophical Review 105:311-35.

Von Mises, Ludwig. 1996. Human Action: A Treatise 011 Economics, 4th ed. Irvington-on-Hudson, N.Y.: Foundation for Economic Education.

Warnock, Mary. 1966. Ethics Since 1900. 2nd ed. London: Oxford University Press.

Weatherson, Brian. 2003. 'What Good arc Counterexamples?' Philosophical Studies 115: 1-31.

Wells, Gary and Elizabeth Olson. 2003. 'Eyewitness Testimony'.

Annual Review of Psychology 54:277-95.

Westermarck, Edward. 1960. Etllical Relativity. Paterson, N.J.: Littlefield, Adams, & Co. Williams, Bernard. 1981. 'Internal and External Reasons'. Pp. 101-13 in Moral LlIck: Philosophical Papers 1973-1980. Cambridge: Cambridge University Press.

Wilson, Edward O. 2000. Sociobiology: The New Synthesis. Cambridge, Mass.: Harvard University Press.

Wright, Crispin. 1992. Truth and Objectivity. Cambridge, Mass.: Harvard University Press.

Wright, Robert. 1995. Tile Moral Animal. New York: Random House.

Zaibert, Leo. 2003. 'In tentions, Promises, and Obligations'. Pp. 52-84 in John Searle, edited by Barry Smith. Cambridge: Cambridge

University Press.